(51) International Patent Classification 6 :		(11) International Publication Number: WO 97/33911
C07K 14/475, 16/26, C12N 15/18, 15/85, 15/63, 1/21, C07H 21/00, A61K 38/18, 48/00, C12Q 1/68, G01N 33/53, C12P 19/34	A1	(43) International Publication Date: 18 September 1997 (18.09.97)

(21) International Application Number: PCT/US97/03461

(22) International Filing Date: 14 March 1997 (14.03.97)

(30) Priority Data:
08/615,944   14 March 1996 (14.03.96)   US

(60) Parent Application or Grant
  (63) Related by Continuation
    US   08/615,944 (CIP)
    Filed on   14 March 1996 (14.03.96)

(71) Applicant (for all designated States except US): WASHING-TON UNIVERSITY [US/US]; One Brookings Drive, St. Louis, MO 63130 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): JOHNSON, Eugene, M., Jr. [US/US]; 13233 Amiot Drive, St. Louis, MO 63146 (US). MILBRANDT, Jeffrey, D. [US/US]; 75 Aberdeen Place, St. Louis, MO 63105 (US). KOTZBAUER, Paul, T. [US/US]; 239 Crandon Drive, St. Louis, MO 63105 (US).

LAMPE, Patricia, A. [US/US]; 10323 Grant Forest Lane, St. Louis, MO 63123 (US).

(74) Agents: HOLLAND, Donald, R. et al.; Howell & Haferkamp, L.C., Suite 1400, 7733 Forsyth Boulevard, St. Louis, MO 63105-1817 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
  With international search report.
  Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PERSEPHIN AND RELATED GROWTH FACTORS

(57) Abstract

A novel growth factor, persephin, which belongs to the GDNF/neurturin family of growth factors, is disclosed. The mouse and rat amino acid sequences have been identified. Mouse and rat persephin genomic DNA sequences have been cloned and sequenced and the respective cDNA sequences identified. In addition, methods for treating degenerative conditions using persephin, methods for detecting persephin gene alterations and methods for detecting and monitoring patient levels of persephin are provided. Methods for identifying additional members of the persephin-neurturin-GDNF family of growth factors are also provided.

PERSEPHIN AND RELATED GROWTH FACTORS

Reference to Government Grant

5 ## Background of the Invention

(1) Field of the Invention

This invention relates generally to trophic or growth factors and, more particularly, to novel growth factors of the neurturin-GDNF family of growth factors.

10 ### (2) Description of the Related Art

The development and maintenance of tissues in complex organisms requires precise control over the processes of cell proliferation, differentiation, survival and function. A major mechanism whereby these

15 processes are controlled is through the actions of polypeptides known as "growth factors". These structurally diverse molecules act through specific cell surface receptors to produce these actions.

Growth factors, termed "neurotrophic factors"

20 promote the differentiation, growth and survival of neurons and reside in the nervous system or in innervated tissues. Nerve growth factor (NGF) was the first neurotrophic factor to be identified and characterized (Levi-Montalcini et al., *J. Exp. Zool. 116*:321, 1951

25 which is incorporated by reference). NGF exists as a non-covalently bound homodimer that promotes the survival and growth of sympathetic, neural crest-derived sensory, and basal forebrain cholinergic neurons. In sympathetic neurons this substance produces neurite outgrowth *in*

30 *vitro* and increased axonal and dendritic growth *in vivo*. (See Levi-Montalcini and Booker, *Proc Nat'l Acad Sci 46*:384-391, 1960; Johnson et al. *Science 210*: 916-918, 1980; Crowley et al., *Cell 76*:1001-12, 1994 which are

2

incorporated by reference). NGF has effects on cognition
and neuronal plasticity, and can promote the survival of
neurons that have suffered damage due to a variety of
mechanical, chemical, viral, and immunological insults

5   (Snider and Johnson, *Ann Neurol* 26:489-506, 1989; Hefti,
J Neurobiol 25:1418-35, 1994 which are incorporated by
reference). NGF also is known to extensively interact
with the endocrine system and in immune and inflammatory
processes. (Reviewed in Scully and Otten, *Cell Biol Int*

10 *19*:459-469, 1995; Otten and Gadient, *Int. J. Devl
Neurosci* 13:147-151, 1995 which are incorporated by
reference). For example, NGF promotes the survival of
mast cells. (Horigome et al. *J Biol Chem* 269:2695-2707,
1994 which is incorporated by reference).

15   In recent years it has become apparent that growth
factors fall into classes, i.e. families or superfamilies
based upon the similarities in their amino acid
sequences. These families include, for example, the
fibroblast growth factor family, the neurotrophin family

20 and the transforming growth factor-beta (TGF-ß) family.
As an example of family member sequence similarities,
TGF-ß family members have 7 canonical framework cysteine
residues which identify members of this superfamily.

          NGF is the prototype of such a family of growth

25 factors. Brain-derived neurotrophic factor (BDNF), the
second member of this family to be discovered, was shown
to be related to NGF by virtue of the conservation of all
six cysteines that form the three internal disulfides of
the NGF monomer (Barde, *Prog Growth Factor Res* 2:237-248,

30 1990 and Liebrock et al. *Nature* 341:149-152, 1989 which
are incorporated by reference). By utilizing the
information provided by BDNF of the highly conserved
portions of two factors, additional members (NT-3, NT-
4/5) of this neurotrophin family were rapidly found by

35 several groups (Klein, *FASEB J* 8:738-44, 1994 which is
incorporated by reference).

3

Neurotrophic factors structurally unrelated to NGF have been recently identified. These include factors originally isolated based upon a "neurotrophic action" such as ciliary neurotrophic factor (CNTF) (Lin et al.,

5 *Science 246*:1023-5, 1989 which is incorporated by reference) along with others originally isolated as a result of non-neuronal activities (e.g. fibroblast growth factors (Cheng and Mattson *Neuron 1*:1031-41,1991 which is incorporated by reference), IGF-I (Kanje et al, *Brain Res*

10 *486*:396-398, 1989 which is incorporated by reference) leukemia inhibitory factor (Kotzbauer et al, *Neuron 12*:763-773, 1994 which is incorporated by reference).

Glial-derived neurotrophic factor (GDNF), is one such neurotrophic factor structurally unrelated to NGF.

15 GDNF was, thus, a unique factor, which, up until now, was not known to be a member of any subfamily of factors. The discovery, purification and cloning of GDNF resulted from a search for factors crucial to the survival of midbrain dopaminergic neurons, which degenerate in

20 Parkinson's disease. GDNF was purified from rat B49 glial cell conditioned media (Lin et al., *Science 260*:1130-2, 1993 which is incorporated by reference). Sequence analysis revealed it to be a distant member of the TGF-ß superfamily of growth factors, having

25 approximately 20% identity based primarily on the characteristic alignment of the 7 canonical framework cysteine residues (Lin et al., *Science 260*:1130-2, 1993 which is incorporated by reference). Thus, GDNF could possibly have represented a new subfamily within the TGF-

30 ß superfamily.

Recombinant GDNF produced in bacteria specifically promotes the survival and morphological differentiation of dopaminergic neurons (Lin et al., *Science 260*:1130-2, 1993); Tomac et al., *Nature 373*:335-9, 1995; Beck et al.,

35 *Nature 373*:339-41, 1995 and Ebendal et al., *J Neurosci Res 40*:276-84, 1995 which are incorporated by reference)

4

and motor neurons (Henderson et al., *Science* 266:1062-4, 1994; Yan et al., *Nature* 373:341-4, 1995; and Oppenheim et al., *Nature* 373:344-6, 1995 which are incorporated by reference). Overall, GDNF was a more potent factor for
5   promoting the survival of motor neurons than the other factors, and it was the only factor that prevented neuronal atrophy in response to these lesions, thereby positioning it as a promising therapeutic agent for motor neuron diseases.
10   It is now generally believed that neurotrophic factors regulate many aspects of neuronal function, including survival and development in fetal life, and structural integrity and plasticity in adulthood. Since both acute nervous system injuries as well as chronic
15   neurodegenerative diseases are characterized by structural damage and, possibly, by disease-induced apoptosis, it is likely that neurotrophic factors play some role in these afflictions. Indeed, a considerable body of evidence suggests that neurotrophic factors may
20   be valuable therapeutic agents for treatment of these neurodegenerative conditions, which are perhaps the most socially and economically destructive diseases now afflicting our society. Nevertheless, because different neurotrophic factors can potentially act preferentially
25   through different receptors and on different neuronal or non-neuronal cell types, there remains a continuing need for the identification of new members of neurotrophic factor families for use in the diagnosis and treatment of a variety of acute and chronic diseases of the nervous
30   system.

Summary of the Invention:
        Briefly, therefore, the present invention is directed to the identification and isolation of substantially purified factors that promote the survival
35   and growth of neurons as well as non-neuronal cells. Accordingly, the inventors herein have succeeded in

5

discovering novel protein growth factors belonging to a
family of growth factors for which GDNF was the first
known member. The first such newly discovered family
member was neurturin and this is the subject of copending
5   application serial Number 08/519,777. Based upon the
sequence of GDNF and neurturin the inventors herein have
discovered another member of the GDNF-Neurturin family of
growth factors referenced herein as persephin (PSP).
This growth factor is believed to show at least 85%
10 sequence identity among homologous sequences from
different mammalian species although sequence homology
may be as low as 65% in non-mammalian species such as
avian species. Persephin proteins identified herein
include mouse sequences as set forth in SEQ ID NOS:79, 80
15 and 81 (Figure 11; amino acid residues 52 through 140, 47
through 142, and 9 through 142, respectively) and rat
sequences as set forth in SEQ ID NOS:82 and 83 (Figure
14; amino acid residues 1 through 89 and 1 through 91,
respectively). In addition, human persephin is
20 identified by virtue of its having at least 85% sequence
homology with its ortholog, mature mouse persephin, along
with the identification of certain conserved amino acid
residues contained within human persephin as shown in
Figure 15. Thus, it is believed that human persephin
25 will have 28 amino acids in the aligned sequence between
the first and seventh canonical framework cysteine
residues as set forth in Figure 15 with residues numbered
from the N-terminal end of the family member aligned
sequence being (1) Cys, (3) Leu, (10) Val, (13) Leu, (14)
30 Gly, (15) Leu, (16) Gly, (17) Tyr, (21) Glu, (25) Phe,
(26) Arg, (27) Tyr, (28) Cys, (30) Gly, (32) Cys, (44)
Leu, (47) Leu, (58) Cys, (59) Cys (61) Pro, (66) Asp,
(69) Phe, (70) Leu, (71) Asp, (83) Ser, (84) Ala, (87)
Cys, and (89) Cys.

35   Persephin has been identified and obtained by a
method based upon the conserved regions of the GDNF-

6

Neurturin family discovered by the inventors herein.
Accordingly, a new method has been devised that utilizes
degenerate primers constructed from the sequences of
these conserved regions for use in the polymerase chain

5 reaction procedure. By utilizing this method the mouse
and rat orthologs of the new family member, persephin,
have been identified and obtained.

The present invention thus provides both amino
acid sequences and nucleotide sequences that encode mouse

10 and rat persephin as set forth in the amino acid
sequences of SEQ ID NOS:79-83 and nucleotide sequences of
SEQ ID NOS:84 and 85. Because of the close homology
between the mouse and rat sequences (95% sequence
identity), it is believed that the human persephin

15 sequence will show a high sequence homology to the mouse
and rat sequences.

Expression vectors and stably transformed cells
are also within the scope of this invention. The
transformed cells can be used in a method for producing

20 persephin.

In another embodiment, the present invention
provides a method for preventing or treating neuronal
degeneration comprising administering to a patient in
need thereof a therapeutically effective amount of

25 persephin. A patient may also be treated by implanting
transformed cells which express persephin or a DNA
sequence which encodes persephin into a patient, or cells
cultured and expanded by growth in persephin.

The present invention also provides compositions

30 and methods for detecting persephin. One method is based
upon persephin antibodies and other methods are based
upon detecting mRNA or cDNA or genomic DNA encoding
persephin using recombinant DNA techniques.

Among the several advantages found to be achieved

35 by the present invention, therefore, may be noted the
provision of a new growth factor, persephin, for use in

7

preventing the atrophy, degeneration or death of certain
cells, in particular neurons; the provision of human
persephin by making available the specific sequences of
murine and rat persephin from which the human sequence

5 can be identified and obtained; the provision of other
members of the neurturin-persephin-GDNF family of growth
factors by making available new methods capable of
obtaining other family members; the provision of methods
for obtaining persephin by recombinant techniques; the

10 provision of methods for preventing or treating diseases
producing cellular degeneration and, particularly
neuronal degeneration; the provision of methods that can
detect and monitor persephin levels in a patient; and the
provision of methods that can detect alterations in the

15 persephin gene.

Brief Description of the Drawings

Figure 1 illustrates the purification scheme for
preparing neurturin from CHO cells;

Figure 2 illustrates the characterization of

20 fractions eluted from Mono S column in purifying
neurturin showing (a) electrophoresis of each fraction on
a SDS-polyacrylamide gel and visualization of the
proteins by silver stain and (b) the neurotrophic
activity present in each fraction in the superior

25 cervical ganglion survival assay;

Figure 3 illustrates the ability of neurturin to
maintain survival of superior cervical ganglionic cells
in culture showing (a) positive control cells maintained
with nerve growth factor (NGF) (b) negative control cells

30 treated with anti-NGF antibodies showing diminished
survival and (c) cells treated with anti-NGF and
neurturin (approximately 3 ng/ml) showing survival of
neurons;

Figure 4 illustrates the concentration-response

35 effect of neurturin in the superior cervical ganglion
survival assay;

8

Figure 5 illustrat s the homology of the amino acid sequences for the mature growth factors, human neurturin (hNTN), mouse neurturin (mNTN), rat GDNF (rGDNF), mouse GDNF (mGDNF) and human GDNF (hGDNF) with
5   identical amino acid residues enclosed in boxes;

Figure 6 illustrates the tissue distribution of neurturin mRNA and the mRNA for GDNF using RT/PCR analysis on RNA samples obtained from embryonic day 21 (E21) and adult rats;
10 Figure 7 illustrates the cDNA and encoded amino acid sequence of human pre-pro neurturin (SEQ ID NO:11) showing the pre- region from nucleic acid 1 through 57 (SEQ ID NO:17), the pro- region from nucleic acid 58 through 285 (SEQ ID NO:20), human neurturin from nucleic
15 acid 286 through 591 (SEQ ID NO:9) and the splice site between nucleic acids 169 and 170 which defines the coding sequence portion of two exons from nucleic acids 1 through 169 (SEQ ID NO:27) and 170 through 594 (SEQ ID NO:28);
20 Figure 8 illustrates the cDNA and encoded amino acid sequence of mouse pre-pro neurturin (SEQ ID NO:12) showing the pre- region from nucleic acid 1 through 57 (SEQ ID NO:18), the pro- region from nucleic acid 58 through 285 (SEQ ID NO:21), mouse neurturin from nucleic
25 acid 286 through 585 (SEQ ID NO:10) and the splice site between nucleic acids 169 and 170 which defines the coding sequence portion of two exons from nucleic acids 1 through 169 (SEQ ID NO:29) and 170 through 588 (SEQ ID NO:30);
30 Figure 9 illustrates the mouse cDNA sequence containing a 5' non-coding region (SEQ ID NO:13) and a 3' non-coding region (SEQ ID NO:14) each of which are contiguous to the coding region of pre-pro neurturin;

Figure 10 illustrates the percent neuronal
35 survival in E18 rat nodose ganglia neurons treated 24 hours post-plating for NTN, GDNF, BDNF, NGF and AMO;

9

Figure 11 illustrates the nucleotide and amino
acid sequence of murine persephin (SEQ ID NOS:79, 80 and
81; amino acid residues 52 through 140, 47 through 142,
and 9 through 142, respectively);

5   Figure 12 illustrates the family member sequence
identity in the region between the first and seventh
canonical framework cysteine residues aligned beginning
with the first canonical framework cysteine for murine
GDNF (SEQ ID NO:87), murine neurturin (NTN) (SEQ ID
10 NO:88) and murine persephin (PSP) (SEQ ID NO:89);

Figure 13 illustrates the partial sequence of rat
persephin cDNA (SEQ ID NO:97) obtained by the technique
of rapid amplification of cDNA ends;

Figure 14 illustrates the partial sequence
15 beginning with the first canonical framework cysteine for
rat persephin (SEQ ID NO:83) and the corresponding
polynucleotide sequence (SEQ ID NO:86);

Figure 15 shows the family member aligned partial
amino acid sequences from the first through the seventh
20 canonical framework cysteine residues illustrating family
member sequence homology of the mature growth factors,
human GDNF, rat GDNF, mouse GDNF, human neurturin (NTN),
mouse neurturin, rat persephin (PSP), and mouse persephin
in which boxes enclose the 28 conserved amino acid
25 residues present in all;

Figure 16 illustrates the sequences of TGF-ß
superfamily members aligned using the Clustal method,
from the first canonical framework cysteine to the end of
the sequence for transforming growth factor-ß1 (TGFß1),
30 transforming growth factor-ß2 (TGFß2), transforming
growth factor-ß3 (TGFß3), inhibin ß A (INHßA), inhibin ß
B (INHßB), the *nodal* gene (NODAL), bone morphogenetic
proteins 2 and 4 (BMP2 and BMP4), the *Drosophila*
decapentaplegic gene (*dpp*), bone morphogenetic proteins
35 5-8 (BMP5, BMP6, BMP7 and BMP8), the *Drosophila* 60A gene
family (60A), bone morphogenetic protein 3 (BMP3), the

10

Vgl g ne, growth differentiation factors 1 and 3 (GDF1
and GDF3), dorsalin (drsln), inhibin α (INHα), the *MIS*
gene (MIS), growth factor 9 (GDF-9), glial-derived
neurotropic growth factor (GDNF) and neurturin (NTN);

5   Figure 17 illustrates full length murine persephin
gene (SEQ ID NO:131), the amino acid sequence which
includes at least a portion of the pre-pro region encoded
by the nucleotide sequence in the first reading frame
from the initiator methionine codon through the stop
10 codon at nucleotide positions 244-246 (SEQ ID NO:132) and
the amino acid sequence which includes mature persephin
in the second reading frame from nucleotide position 2
through the stop codon at positions 557-559 (SEQ ID
NO:133);

15   Figure 18 illustrates full length rat persephin
gene (SEQ ID NO:134), the amino acid sequence which
includes at least a portion of the pre-pro region encoded
by the nucleotide sequence in the first reading frame
from the initiator methionine codon through the stop
20 codon at nucleotide positions 244-246 (SEQ ID NO:135) and
the amino acid sequence which includes mature persephin
in the second reading frame from nucleotide position 2
through the stop codon at positions 557-559 (SEQ ID
NO:136);

25   Figure 19 illustrates a western blot analysis
using anti-persephin antibodies to detect persephin
protein in cell lysates from COS monkey cells transfected
with the murine persephin gene (lane 2) or the rat
persephin gene (lane 3) compared to cells transfected
30 with the non-recombinant vector alone (pCB6, lane 4) and
the mature protein produced by *E. Coli* (lane 1);

          Figure 20 illustrates the murine chimeric
molecules (A) PSP/NTN containing the persephin fragment
(residues 1-63) and the neurturin fragment (residues 68-
35 100) and (B) NTN/PSP containing the neurturin fragment
(residues 1-67) and the persephin fragment (residues 64-

96) with the arrow indicating the crossover point in each;

Figure 21 illustrates the survival promoting effect of persephin in murine embrionic day-14
5 mesencephalic cells cultured for three days (a) in the absence of persephin where almost all of the cells are dead and (b) in the presence of persephin (100 ng/ml) where substantial neuronal cell survival is evident; and

Figure 22 illustrates RT/PCT survey for persephin
10 expression in adult mouse tissues showing persephin expression by Kidney cells.

Description of the Preferred Embodiments

The present invention is based upon the
15 identification, isolation and sequencing of a DNA molecule that encodes a new growth factor, persephin. Because of the sequence similarity to neurturin and GDNF, persephin is believed to be capable of promoting cell survival and, in particular, the survival of neurons.
20 Prior to this invention, persephin was unknown and had not been identified as a discrete biological substance nor had it been isolated in pure form.

The growth factor, neurturin (NTN) was identified and isolated as set forth in copending application Serial
25 Number 08/519,777 filed August 28, 1995, which is incorporated in its entirety by reference. From the sequence of neurturin and the sequence of the closely related growth factor, glial-derived neurotrophic factor (GDNF), the inventors herein have devised and pursued
30 strategies to find additional related factors. Neurturin is approximately 40% identical to GDNF, but less than 20% identical to any other member of the TGF-ß superfamily. Together these two proteins define a new subfamily within the TGF-ß superfamily. Several sequence regions within
35 neurturin and GDNF were identified that are highly conserved, such that they are likely to be present in any

12

additional members of this subfamily. This sequence
information can therefore be used to isolate previously
unknown members of this subfamily by designing degenerate
oligonucleotides to be used as either primers in PCR
5 reactions or as probes in hybridization studies.

Using the new degenerate primer PCR strategy
described in Example 11 of copending application Serial
Number 08/519,777, the inventors herein have succeeded in
identifying a third factor, persephin, that is
10 approximately 40-50% identical to both GDNF and
neurturin. Primers corresponding to the amino acid
sequence from conserved regions of neurturin and GDNF
(SEQ ID NO:42 and SEQ ID NO:44) were used to amplify a 77
nt fragment from rat genomic DNA. The resulting products
15 were subcloned into the Bluescript KS plasmid and
sequenced. The sequence of one of the amplified products
predicted amino acid sequence data internal to the PCR
primers that was different from that of GDNF or neurturin
but had more than 20% identity with GDNF and neurturin,
20 whereas the sequences of other amplified products we
obtained corresponded to GDNF or neurturin, as would be
expected. The 22 nucleotide sequence (SEQ ID NO:90) was
then aligned with the rat sequences of GDNF and neurturin
and found to be unique. This novel sequence, thus,
25 suggested that we had identified a new family member
referenced herein as persephin.

To obtain additional persephin sequence
information, primers containing the unique 22 nucleotide
sequence of the amplified fragment were used in the rapid
30 amplification of cDNA ends (RACE) technique (Frohman,
M.A. Methods in Enzymology 218:340-356, 1993) using cDNA
obtained from neonatal rat brain. An approximately 350
nt fragment was obtained from this PCR reaction which
constituted a partial rat persephin cDNA sequence of
35 approximately 350 nucleotides (SEQ ID NO:106). The
predicted amino acid sequence of this cDNA was compared

13

to that of GDNF and neurturin, and found to have
approximately 40% identity with each of these proteins.
Importantly, the characteristic spacing of the canonical
framework cysteine residues in members of the TGF-β

5 superfamily was present. Furthermore, in addition to the
region of similarity encoded by the degenerate primers
used to isolate persephin, another region of high
homology shared between GDNF and neurturin, but absent in
other members of the TGF-β superfamily, was also present

10 in persephin

GDNF	ACCRPVAFDDDLSFLDD	(aa 60-76)	(SEQ ID NO:98)
NTN	PCCRPTAYEDEVSFKDV	(aa 61-77)	(SEQ ID NO:99)
PSP	PCCQPTSYAD-VTFLDD	(aa 57-72)	(SEQ ID NO:100)

15

(Amino acid numbering uses the first Cys residue as amino
acid 1).
        With the confirmation that persephin was indeed a
new member of the GDNF/NTN subfamily, we isolated murine

20 genomic clones of persephin to obtain additional sequence
information. Primers corresponding to rat cDNA sequence
were used in a PCR reaction to amplify a 155 nucleotide
(nt) fragment from mouse genomic DNA which was homologous
to the rat persephin cDNA sequence. These primers were

25 then used to obtain murine persephin genomic clones from
a mouse 129/Sv library in a P1 bacteriophage vector
(library screening service of Genome Systems, Inc., St.
Louis, MO).
        Restriction fragments (3.4 kb Nco I and a 3.3 kb Bam

30 H1) from this P1 clone containing the persephin gene were
identified by hybridization with a 210 nt fragment of
persephin obtained by PCR using mouse genomic DNA and
persephin-specific primers. The Nco I and Bam H1
fragments were sequenced and found to encode a stretch of

35 amino acids corresponding to that present in the rat
persephin RACE product, as well as being homologous to
the mature regions of both neurturin and GDNF (Figure
11).
        When the amino acid sequences of mature murine GDNF,

14

NTN and PSP are aligned using the first canonical framework cysteine as the starting point, which is done because alterations in the cleavage sites between family members creates variability in the segments upstream of

5   the first cysteine, persephin (91 amino acids) is somewhat smaller than either neurturin (95 amino acids) or GDNF (94 amino acids). The overall identity within this region is about 50% with neurturin and about 40% with GDNF (Figure 12).

10   Further nucleotide sequencing of the murine persephin NcoI fragment revealed the nucleotide sequence of the entire murine persephin gene as shown in Figure 17. In addition, the entire rat persephin gene has been determined by sequencing a PCR amplified fragment of rat

15 genomic DNA as shown in Figure 18. In both the murine and rat persephin gene, an open reading frame extends from the sequence coding for an initiator methionine up to a stop codon at positions 244-246. However, somewhere in this sequence an apparent anomaly occurs such that the

20 sequence encoding the RXXR cleavage site (positions 257-268) and the sequence corresponding to the mature persephin protein (positions 269-556) are not co-linear with this open reading frame. Instead, a second reading frame encodes the cleavage site and the mature persephin.

25 The two cogent reading frames are shown in Figures 17 and 18. Irrespective of this apparent anomaly, mammalian cells express persephin from either the murine or rat full length genomic sequence (see Example 14 below).
        The N-terminus of persephin was predicted by

30 reference to the N-terminal regions of neurturin or GDNF. Using neurturin sequence homology and cleavage signals, a characteristic RXXR cleavage motif is present beginning 9 residues upstream of the first canonical framework cysteine of persephin which would suggest that mature

35 murine persephin would contain 5 amino acids (ALAGS) (SEQ ID NO:103) upstream of this cysteine (as does neurturin).

15

The corresponding 5 amino acids in rat persephin are ALPGL (SEQ ID NO:112). Using these parameters, mature persephin would consist of 96 amino acids and have a predicted molecular mass of 10.4 kD. Using GDNF sequence

5   homology and cleavage signals, on the other hand suggests that the N-terminus upstream from the first cysteine of persephin could be longer, in accord with that observed for GDNF which is 40 residues. A characteristic RXXR cleavage motif is thus located 47 residues upstream of

10 the first cysteine and this would suggest that mature persephin would contain 43 amino acids (VRIPGGLPTPQFLLSKPSLCLTILLYLALGNNHVRLPRALAGS) (SEQ ID NO:104) upstream of this cysteine. Using these parameters, mature persephin would consist of 134 amino

15 acids and have a predicted molecular mass of 14.5 kD. Thus, mature persephin may exist in either or both of the 96 amino acid predicted 10.4 kD or for the 134 amino acid predicted 14.5 kD form.

By "mature" growth factor reference is made to the

20 secreted form of the growth factor in which any pre- or pro- regions have been cleaved and which may exist as a monomer or, by analogy to other members of the TGF-ß superfamily, in the form of a homodimer linked by disulfide bonds.

25   The discovery of the new growth factor, persephin, as described above is a result of the prior discovery by the inventors herein of neurturin. Thus, the experiments leading to the discovery of neurturin are relevant to the current discovery of persephin and to the predicted human

30 form of persephin as well as to the biological activity of persephin.

Neurturin was identified and isolated by the inventors herein from conditioned medium for CHO cells. The initial neuronal survival promoting activity was

35 identified by the inventors in a partially purified preparation of this CHO-conditioned medium. Preparation

16

of conditioned medium for a given cell line is well known
in the art (for example, see Reid, in *Methods in
Enzymology Vol. LVIII, Cell Culture*, Jakoby and Pastan,
Eds., Academic Press, San Diego, pp 161-164, 1979;
5   Freshney, Culture of Animal Cells in *A Manual of Basic
Technique*, 2d Ed., Wiley-Liss, NY, p. 84, 1987 which are
incorporated by reference). Thus, although in the
present work CHO cells were cultured and the conditioned
medium used to identify and to obtain neurturin in
10 purified form, one skilled in the art will readily
appreciate that any cell that expresses neurturin can be
used as a source. Some of the cells that express
neurturin are identified below in Example 9 and the
inventors herein believe that any of the cells identified
15 as expressing neurturin can be used to obtain conditioned
medium from which neurturin can be isolated.

In the isolation of neurturin from the CHO cell
conditioned medium, an initial crude conditioned medium
can be obtained by centrifugation and/or filtration to
20 remove cellular debris. For further purification, one
skilled in the art will readily appreciate that any of a
number of methods known in the art can be used to isolate
and purify neurturin from a biological sample such as
affinity chromatography, ion exchange chromatography,
25 preparative electrophoresis or the like wherein the
methods are used either individually or in combination.

The cell survival promoting effect of neurturin can
be assessed in any suitable system for assessing cell
survival. The inventors herein believe that neurturin
30 can promote survival in a variety of different tissues
based upon what is known for other growth factors and
upon the observation that neurturin is expressed in a
number of tissues in which it is believed to have a
survival promoting effect.
35   By virtue of the degree of sequence identity of
persephin with its paralogs, neurturin and GDNF and the

17

known actions of these substances in promoting the
survival and growth of neuronal and non-neuronal tissues,
it is also believed that persephin will promote survival
and growth in neuronal tissues as well as a variety of
5   non-neuronal tissues. Indeed, the inventors herein have
identified brain, kidney, and heart tissues as tissues.
expressing persephin which further supports the
conclusion that persephin can act to promote growth and
survival in neuronal and non-neuronal cells.
10   In the work reported herein, neuronal activity for
neurturin was assessed using a sympathetic neuronal
survival assay (sympathetic cervical ganglia, SCG) which
has been extensively characterized (Martin et al, *J Cell
Biol 106*:829-844, 1989; Deckwerth and Johnson, *J Cell
15   Biol 123*:1207-1222, 1993 which are incorporated by
reference) (see Figure 3). We also show the survival
promoting effects of neurturin on sensory neurons (See
Figure 10). In the same sympathetic and sensory neuronal
cell assays, persephin showed little survival
20   potentiating activity. However, in a preparation of CNS
neuronal cells of mesencephalic origin, persephin showed
neuronal survival potentiating activity. This suggests
that persephin will be applicable in the treatment or
prevention of diseases involving neuronal degeneration in
25   the CNS such as, for example, Parkinson's disease.
         By way of illustrating the methods used in the above
survival assays, the SCG assay involved the culturing of
cells obtained from superior cervical ganglia of rat
embryo for 5 days at 37°C in medium containing nerve
30   growth factor (NGF). The medium was then exchanged with
a medium containing no NGF and containing anti-NGF
antiserum. Removal of NGF results normally in death of
the neurons in 24-72 hours. Neuronal survival was
visually assessed under a microscope on days 7-8.
35   Maximum neuronal survival criteria included lack of
degeneration of both neuronal cell bodies and neurites.

Cell body degeneration was indicated when the neuronal
cell body was reduced in size, showed irregular membrane
swellings, contained vacuoles, or had lost refractility.
A field of neurites was scored as showing signs of
5 disintegration when swellings and blebs appeared along
the neurite bundles. Survival was determined by
comparison with neurons grown in the presence of NGF
(positive control) or in the absence of NGF with NGF
antisera (negative control).

10 Activity was quantitated by calculation of a
"survival unit". The total survival units in a sample
were defined as the minimal volume of an aliquot of the
sample which produced maximal survival divided into the
total volume of that sample. For example, a volume of
15 600 ml was eluted from the heparin agarose column and
from this eluate, 12.5 μl was the minimum volume that
promoted maximal volume. Thus, the survival units in the
eluate from the heparin agarose column was 48,000.
Specific activity was calculated as the survival units
20 divided by the mg total protein. The intrinsic activity
of neurturin is expressed herein in concentration units
of pg/ml or pM promoting maximal or half-maximal
survival. As shown in Figure 5, a concentration-response
curve of purified neurturin protein indicates that the
25 intrinsic activity of neurturin expressed as an EC_{50} is
approximately 1.5 ng/ml or approximately 50 pM and an
EC_{100} is approximately 3 ng/ml or approximately 100 pM.

Survival units were determined in an assay using
approximately 1200 neurons in a 0.5 ml culture assay and
30 a culture period of 48 hours following addition of the
fraction. Survival was assessed visually after the 48
hours. Intrinsic activity as shown in Figure 4 was
determined in an assay using approximately 2700 neurons
and a culture period of 72 hours. Survival was assessed
35 by fixing the neurons and counting the number of
surviving neurons. Because the stability, as assessed by

19

half-life of activity, for neurturin decreases as the
number of neurons increases, the intrinsic activity
measurement would be expected to be lower than that
predicted by Specific Activity determinations. The

5   intrinsic activity measurement would also be expected to
be lower than that predicted by specific activity because
the survival was measured after 72 hours instead of 48
hours.

     The purification of neurturin is described in detail
10   in Example 1 below. The conditioned medium starting
material was prepared from a derivative of DG44 Chinese
hamster ovary cells, DG44CHO-pHSP-NGFI-B (Day et al, J
Biol Chem 265:15253-15260, 1990 which is incorporated by
reference). The inventors herein have also isolated

15   neurturin in partially purified form from conditioned
medium of other derivatives of DG44 Chinese hamster ovary
cells and these other cells could be used equally as well
as the DG44CHO-pHSP-NGFI-B cells as could the parent DG44
Chinese hamster ovary Cells, ovary cells from other

20   species and cells from other tissues such as those known
to express neurturin (See example 9). In preparing the
conditioned medium, cells were placed in serum free
medium for 2 days at which time conditioned medium is
collected and the medium replenished. This cycle was

25   repeated to yield 5 harvests of conditioned medium from
each batch of CHO cells. The collected media was
centrifuged to remove cellular debris.

     The first step in purification of neurturin from the
CHO cell conditioned medium involved the introduction of

30   the conditioned medium onto a heparin agarose column and
the elution of partially purified neurturin therefrom.
This step resulted in an 111 fold increase in the
specific activity and purification of the protein. The
buffer used to apply the medium to the column contains

35   0.5 M NaCl. At this concentration of NaCl the neurturin
binds to the heparin agarose matrix. The inventors

20

herein believe that based upon their isoelectric points,
LIF and CNTF would either not bind to the heparin agarose
matrix or be washed away from the matrix with buffer
containing 0.5 M NaCl. Thus, this step would be expected
5 to isolate neurturin from growth factors such as LIF and
CNTF. After washing the column, neurturin was eluted
from the column using 1.0 M NaCl.

     For further purification, the eluted material was
then diluted and introduced into a column containing SP
10 SEPHAROSE® High Performance ion exchange resin
(Pharmacia, Piscataway, NJ). Material eluted from this
column was further purified using fast protein liquid
chromatography (FPLC) on a Chelating Superose HR 10/2
column charged with Cu^{++} (Pharmacia, Piscataway, NJ).
15 Eluted fractions from the Cu^{++} superose column were
introduced into a Mono S HR 5/5 cation exchange column
(Pharmacia, Piscataway, NJ) for further FPLC
purification. The composition of the proteins in the
Mono S fractions were analyzed using non-reducing SDS-
20 PAGE and silver staining.

     Fractions collected from the columns at each stage
of purification were assayed for biological activity
using the neuronal survival assay and for protein content
using the dye binding method of Bradford (*Anal Biochem*
25 *72*:248-254, 1976 which is incorporated by reference) with
a Bio-Rad protein assay dye reagent (Bio-Rad
Laboratories, Inc., Hercules, CA). The progressive
purification using the above steps is shown in table 1.

21

Table 1

	Protein[a] (mg)	Activity[b] (units)	Specific Activity[d] (units/mg)	Yield (%)	Purification (fold)
Conditioned Medium	5000	48000[c]	9.6	–	–
Heparin Agarose	45	48000	1068	100	111
SP Sepharose	5.3	48000	9058	100	943
Cu++ Superose	0.31	30000	96700	62	10070
Mono S	0.004	15000	3750000	31	390000

a. mg protein was determined using the dye binding method of Bradford (*Anal Biochem* 72:248, 1976).

b. The total activity units or survival units in a sample were defined as the minimal volume of an aliquot of the sample which produced maximal survival divided into the total volume of that sample.

c. Activity for Conditioned Medium was derived from the assumption that 100% of the activity was recovered in the heparin agarose fraction because the activity of conditioned medium was too low to be directly assayed.

d. Specific Activity was the Activity units divided by the mg total protein.

The results of this analysis along with the results of the neuronal survival assay of fractions revealed that a protein having an apparent molecular weight of about 25 kD co-purified with the sympathetic neuron survival activity.

The purified material isolated from CHO cell conditioned medium was used to determine partial amino acid sequences of the protein in CHO cell conditioned medium and subsequently as a basis for determining the sequences in different species. The N-terminal amino acid sequence was determined using an automated protein/peptide sequencer and the first 16 amino acids

22

were considered to be, with uncertainty as to position 6,
Ser-Gly-Ala-Arg-Pro-Xaa-Gly-Leu-Arg-Glu-Leu-Glu-Val-Ser-
Val-Ser where Xaa was an unknown amino acid (SEQ ID
NO:3). Internal amino acid fragments were obtained from
5   the purified material following digestion with protease
enzymes and the sequences determined. Three internal
fragments thus obtained were (1) with uncertainty as to
positions 1, 2 and 6, Xaa₁-Cys-Ala-Gly-Ala-Xaa₂-Glu-Ala-
Ala-Val where Xaa₁ was unknown amino acid, Xaa₂ was Ser or
10   Cys (SEQ ID NO:4); (2) with uncertainty as to positions
1, 2, 4, 10, 17 and 22, Xaa₁-Xaa₂-Val-Glu-Ala-Lys-Pro-Cys-
Cys-Gly-Pro-Thr-Ala-Tyr-Glu-Asp-Xaa₃-Val-Ser-Phe-Leu-Ser-
Val where Xaa₁ and Xaa₂ were unknown, Xaa₃ was Gln or Glu
(SEQ ID NO:5) and (3) Tyr-His-Thr-Leu-Gln-Glu-Leu-Ser-
15   Ala-Arg (SEQ ID NO:6). Based upon these partial amino
acid sequences, DNA probes and primers can be made and
used to obtain cDNA clones from different species based
upon high sequence conservation between mammalian
species. The human cDNA and inferred amino acid sequence
20   is shown in Figure 7 and the mouse cDNA and inferred
amino acid sequence is shown in Figure 8.

The cDNA clone from mouse was 1.0 kb having an
open reading frame of 585 nucleotides (SEQ ID NO:12)
encoding the mouse pre-pro neurturin protein (SEQ ID
25   NO:8, Figure 8). In addition, non-coding regions have
been identified at both the 5' and 3' ends of the coding
region as shown in Figure 9. (SEQ ID NO:13, 5' non-coding
region, nucleic acids -348 through -1; SEQ ID NO:14, 3'
non-coding region, nucleic acids 589 through 675). The
30   mouse neurturin sequence can be used to obtain PCR
primers for use in identifying homologs from other
species. A human 192 nucleotide fragment from human
genomic DNA was amplified by this method and further used
to screen a human genomic library to obtain clones
35   containing the human neurturin genomic locus. The human
cDNA sequence was deduced from the sequencing of these

23

clones. (Figure 7, cDNA sequence of human pre-pro
neurturin).

Reference to persephin or to neurturin herein is
intended to be construed to include growth factors of any
5   origin which are substantially homologous to and which
are biologically equivalent, respectively, to the
persephin characterized and described herein or to the
neurturin characterized and described herein. Such
substantially homologous growth factors may be native to
10   any tissue or species and, similarly, biological activity
can be characterized in any of a number of biological
assay systems. Reference to pre-pro neurturin herein is
intended to be construed to include pre-pro growth
factors containing a pre- or leader or signal sequence
15   region, a pro- sequence region and neurturin as defined
herein.

The terms "biologically equivalent" are intended
to mean that the compositions of the present invention
are capable of demonstrating some or all of the same
20   growth properties in a similar fashion, not necessarily
to the same degree as the neurturin isolated from the CHO
cell conditioned medium herein or recombinantly produced
human or mouse or rat neurturin or persephin as the case
may be.

25   By "substantially homologous" it is meant that the
degree of sequence identity of neurturin orthologs
including human and mouse neurturin as well as neurturin
from any other species or the degree of sequence identity
of persephin orthologs including human, mouse and rat
30   persephin as well as persephin from any other species, is
greater than that between paralogs such as persephin and
neurturin or persephin and GDNF, and greater than that
reported previously for members of the TGF-ß superfamily
(For discussion of homology of TGF-ß superfamily members
35   see Kingsley, *Genes and Dev* 8:133-46, 1994 which is
incorporated by reference).

24

Sequence identity or percent identity is intended
to mean the percentage of same residues between two
sequences. The reference sequence is mouse persephin
when determining percent identity with mouse GDNF and
5 mouse neurturin and rat persephin when determining
percent identity with rat GDNF and rat neurturin.
Referencing is to human neurturin when determining
percent identity with non-human neurturin, to human
neurturin when determining percent identity with non-
10 neurturin growth factors and to human GDNF when
determining percent identity of non-neurturin growth
factors with GDNF. In all of the above comparisons, the
two sequences being compared are aligned using the
Clustal method (Higgins et al, *Cabios 8*:189-191, 1992) of
15 multiple sequence alignment in the Lasergene biocomputing
software (DNASTAR, INC, Madison, WI). In this method,
multiple alignments are carried out in a progressive
manner, in which larger and larger alignment groups are
assembled using similarity scores calculated from a
20 series of pairwise alignments. Optimal sequence
alignments are obtained by finding the maximum alignment
score, which is the average of all scores between the
separate residues in the alignment, determined from a
residue weight table representing the probability of a
25 given amino acid change occurring in two related
proteins over a given evolutionary interval. Penalties
for opening and lengthening gaps in the alignment
contribute to the score. The default parameters used
with this program are as follows: gap penalty for
30 multiple alignment = 10; gap length penalty for multiple
alignment = 10; k-tuple value in pairwise alignment = 1;
gap penalty in pairwise alignment = 3; window value in
pairwise alignment = 5; diagonals saved in pairwise
alignment = 5. The residue weight table used for the
35 alignment program is PAM250 (Dayhoff et al., in *Atlas of*

25

Protein Sequence and Structure, Dayhoff, Ed., NBRF,
Washington, Vol. 5, suppl. 3, p. 345, 1978).

Percent conservation is calculated from the above
alignment by adding the percentage of identical residues
5 to the percentage of positions at which the two residues
represent a conservative substitution (defined as having
a log odds value of greater than or equal to 0.3 in the
PAM250 residue weight table). Conservation is referenced
to mouse persephin when determining percent conservation
10 with persephin from other species or with non-persephin
growth factors; referenced to human neurturin when
determining percent conservation with non-human neurturin
or with non-neurturin growth factors, and referenced to
human GDNF when determining percent conservation to non-
15 persephin, non-neurturin growth factors with GDNF.
Conservative amino acid changes satisfying this
requirement are: R-K; E-D, Y-F, L-M; V-I, Q-H.

Table 2 shows the calculations of identity (I) and
conservation (C) for comparisons of persephin and
20 neurturin and GDNF from various species. Comparisons
were made between mouse persephin from the first
canonical framework cysteine to the end (SEQ ID NO:89)
and rat persephin from the first canonical framework
cysteine to the end (SEQ ID NO:83); between mouse
25 persephin and mouse GDNF from the first cysteine to the
end (mGDNF/C-END, SEQ ID NO:87) or mouse neurturin from
the first cysteine to the end (mNTN/C-END, SEQ ID NO:88);
and between rat persephin and rat GDNF from the first
cysteine to the end (rGDNF/C-END). Neurturin comparisons
30 were between mature human and mature mouse neurturin
(hNTN and mNTN, respectively) and between each of these
and mature human, rat and mouse GDNF (hGDNF, rGDNF and
mGDNF, respectively) as shown in the table.

26

Table 2.

	COMPARISON	% IDENTITY	% CONSERVATION
5	mPSP v. rPSP	96	98
	mPSP v. mNTN/C-END	51	54
	mPSP v. mGDNF/C-END	41	46
	rPSP v. rGDNF/C-END	42	45
10	hNTN v. mNTN	90	93
	hNTN v. rGDNF	44	53
	hNTN v. mGDNF	43	52
	hNTN v. hGDNF	43	53
	mNTN v. rGDNF	42	52
15	mNTN v. mGDNF	41	51
	mNTN v. hGDNF	41	52

    The degree of homology between the mouse persephin

20 and rat persephin is about 96% and it is believed that
    the degree of homology between either mouse or rat
    persephin and human persephin is at least about 85%
    identity based upon a similar comparison with neurturin.
    The neurturin comparisons as shown in Table 2 indicate

25 mature mouse and human neurturin proteins have about 90%
    sequence identity. Furthermore, all persephin and
    neurturin homologs of non-human mammalian species are
    believed to similarly have at least about 85% sequence
    identity with human persephin and neurturin,

30 respectively. For non-mammalian species such as avian
    species, it is believed that the degree of homology with
    persephin or neurturin is at least about 65% identity
    with human persephin or human neurturin, respectively.
    By way of comparison, the variations between family

35 members of the neurturin-persephin-GDNF family of growth
    factors can be seen by the comparison of persephin and
    GDNF or neurturin and GDNF. Mouse and rat persephin have
    about 35 to 40% sequence identity with mouse and rat GDNF
    respectively. Similarly, human and mouse neurturin have

40 about 40% sequence identity and about 50% sequence
    conservation with human, mouse and rat GDNF. It is
    believed that the different family members also have a

27

similar sequence identity of about 40% of that of
neurturin, about 40% of that of persephin and about 40%
of that of GDNF and within a range of about 30% to about
85% identity with neurturin, within a range of about 30%
5   to about 85% identity with persephin and within a range
of about 30% to about 85% sequence identity with GDNF.
Thus, a given member of the GDNF-neurturin-persephin
family would be expected to have lesser sequence identity
with any other family member of the same species than is
10 present in orthologs of that family member in other
species just as human GDNF and human neurturin are more
closely related to mouse GDNF and mouse neurturin,
respectively, than to each other or to GDNF and any given
family member would be expected to have greater sequence
15 identity with another family member than to any other
known member of the TGF-β superfamily (Kingsley, supra).
        In the case of pre-pro neurturin, homologs of pre-
pro neurturin in non-human mammalian species can be
identified by virtue of the neurturin portion of the
20 amino acid sequence having at least about 85% sequence
identity with human neurturin and homologs of pre-pro
neurturin in non-mammalian species can be identified by
virtue of the neurturin portion of the amino acid
sequence having at least about 65% identity with human
25 neurturin. It is similarly believed that mammalian pre-
pro persephin proteins including the human ortholog will
have at least about 85% sequence identity in the mature
persephin portion of the molecule and that non-mammalian
pre-pro persephin proteins will have at least about 65%
30 sequence identity with human pre-pro persephin.
        Either persephin or neurturin, as the terms are used
herein, can also include hybrid and modified forms of
persephin or neurturin, respectively, including fusion
proteins and persephin or neurturin fragments and hybrid
35 and modified forms in which certain amino acids have been
deleted or replaced and modifications such as where one

or more amino acids have been changed to a modified amino
acid or unusual amino acid and modifications such as
glycosolations so long as the hybrid or modified form
retains the biological activity of persephin or
5   neurturin. By retaining the biological activity, it is
meant that neuronal survival is promoted, although not
necessarily at the same level of potency as that of the
neurturin isolated from CHO cell conditioned medium or
that of the recombinantly produced human or mouse
10 neurturin or human or mouse or rat persephin.

Also included within the meaning of substantially
homologous is any persephin or neurturin which may be
isolated by virtue of cross-reactivity with antibodies to
the persephin or neurturin, respectively, as described
15 herein or whose encoding nucleotide sequences including
genomic DNA, mRNA or cDNA may be isolated through
hybridization with the complementary sequence of genomic
or subgenomic nucleotide sequences or cDNA of the
persephin or neurturin, respectively, as described herein
20 or fragments thereof. It will also be appreciated by one
skilled in the art that degenerate DNA sequences can
encode human neurturin or human persephin and these are
also intended to be included within the present invention
as are allelic variants of neurturin or persephin,
25 respectively.

In the case of pre-pro neurturin, alternatively
spliced protein products resulting from an intron located
in the coding sequence of the pro region may exist. The
intron is believed to exist in the genomic sequence at a
30 position corresponding to that between nucleic acids 169
and 170 of the cDNA which, in turn, corresponds to a
position within amino acid 57 in both the mouse and human
pre-pro neurturin sequences (see Figures 7 and 8). Thus,
alternative splicing at this position might produce a
35 sequence that differs from that identified herein for
human and mouse pre-pro neurturin (SEQ ID NO:11 and SEQ

29

ID NO:12, respectively) at the identified amino acid site
by addition and/or deletion of one or more amino acids.
Any and all alternatively spliced pre-pro neurturin
proteins are intended to be included within the terms

5 pre-pro neurturin as used herein.

   Although it is not intended that the inventors,
herein be bound by any theory, it is thought that the
human and mouse proteins identified herein as well as
homologs from other tissues and species may exist as

10 dimers in their biologically active form in a manner
consistent with what is known for other factors of the
TGF-ß superfamily.

   In addition to homodimers, the monomeric units of
the dimers of neurturin or persephin can be used to

15 construct stable growth factor heterodimers or
heteromultimers comprising at least one monomer unit
derived from persephin or at least one monomer unit
derived from neurturin. This can be done by dissociating
a homodimer of neurturin or a homodimer of persephin into

20 its component monomeric units and reassociating in the
presence of a monomeric unit of a second or subsequent
homodimeric growth factor. This second or subsequent
homodimeric growth factor can be selected from a variety
of growth factors including neurturin, persephin, GDNF, a

25 member of the NGF family such as NGF, BDNF, NT-3 and NT-
4/5, a member of the TGF-ß superfamily, a vascular
endothelial growth factor, a member of the CNTF/LIF
family and the like.

   Growth factors are thought to act at specific

30 receptors. For example, the receptors for TGF-ß and
activins have been identified and make up a family of
Ser/Thr kinase transmembrane proteins (Kingsley, *Genes
and Dev* 8:133-146, 1994; Bexk et al *Nature 373*:339-341,
1995 which are incorporated by reference). In the NGF

35 family, NGF binds to the TrkA receptor in peripheral
sensory and sympathetic neurons and in basal forebrain

30

neurons; BDNF and NT-4/5 bind to trkB receptors; and NT-3
binds primarily to trkC receptors that possess a distinct
distribution within the CNS (Tuszynski et al., *Ann Neurol*
35:S9-S12, 1994). The inventors herein believe that
5   persephin, neurturin, GDNF and as yet unknown members of
this family of growth factors act through specific
receptors having distinct distributions as has been shown
for other growth factor families. These may be separate
receptors or it is also possible that members of the
10 GDNF-neurturin-persephin family may act upon the same
receptor as is the case with BDNF and NT-4/5 which act on
the trkB receptor. Nevertheless, by forming heterodimers
or heteromultimers of persephin or neurturin and one or
more other growth factors, the resultant growth factor
15 would be expected to be able to bind to at least two
distinct receptor types preferentially having a different
tissue distribution. The resultant heterodimers or
heteromultimers would be expected to show an enlarged
spectrum of cells upon which it could act or provide
20 greater potency. It is also possible that the
heterodimer or heteromultimer might provide synergistic
effects not seen with homodimers or homomultimers. For
example, the combination of factors from different
classes has been shown to promote long-term survival of
25 oligodendrocytes whereas single factors or combinations
of factors within the same class promoted short-term
survival (Barres et al., *Development* 118:283-295, 1993).
       Heterodimers can be formed by a number of methods.
For example, homodimers can be mixed and subjected to
30 conditions in which dissociation/unfolding occurs, such
as in the presence of a dissociation/unfolding agent,
followed by subjection to conditions which allow monomer
reassociation and formation of heterodimers.
Dissociation/unfolding agents include any agent known to
35 promote the dissociation of proteins. Such agents
include, but are not limited to, guanidine hydrochloride,

31

urea, potassium thiocyanate, pH lowering agents such as
buffered HCl solutions, and polar, water miscible organic
solvents such as acetonitrile or alcohols such as
propanol or isopropanol. In addition, for homodimers
5 linked covalently by disulfide bonds as is the case with
TGF-ß family members, reducing agents such as
dithiothreitol and ß-mercaptoethanol can be used for
dissociation/unfolding and for reassociation/refolding.

Heterodimers can also be made by transfecting a cell
10 with two or more factors such that the transformed cell
produces heterodimers as has been done with the
neurotrophins. (Heymach and Schooter, *J Biol Chem*
270:12297-12304, 1995).

Another method of forming heterodimers is by
15 combining persephin or neurturin homodimers and a
homodimer from a second growth factor and incubating the
mixture at 37°C.

When heterodimers are produced from homodimers, the
heterodimers may then be separated from homodimers using
20 methods available to those skilled in the art such as,
for example, by elution from preparative, non-denaturing
polyacrylamide gels. Alternatively, heterodimers may be
purified using high pressure cation exchange
chromatography such as with a Mono S cation exchange
25 column or by sequential immunoaffinity columns.

It is well known in the art that many proteins are
synthesized within a cell with a signal sequence at the
N-terminus of the mature protein sequence and the protein
carrying such a leader sequence is referred to as a
30 preprotein. The pre- portion of the protein is cleaved
during cellular processing of the protein. In addition
to a pre- leader sequence, many proteins contain a
distinct pro sequence that describes a region on a
protein that is a stable precursor of the mature protein.
35 Proteins synthesized with both pre- and pro- regions are
referred to as preproteins. In view of the processing

32

events known to occur with other TGF-β family members as
well as the sequences determined herein, the inventors
believe that the forms of persephin or neurturin protein
as synthesized within a cell is the pre-pro persephin or
5 a pre-pro neurturin. In the case of neurturin, the pre-
pro neurturin is believed to contain an N-terminal 19
amino acid signal sequence (human pre- signal sequence,
SEQ ID NO:15, Figure 7, amino acids 1 through 19 encoded
by SEQ ID NO:17, Figure 7, nucleic acids 1 through 57;
10 mouse pre- signal sequence, SEQ ID NO:16, Figure 8, amino
acids 1 through 19, encoded by SEQ ID NO:18, Figure 8,
nucleic acids 1 through 57). It is known that the full
length of a leader sequence is not necessarily required
for the sequence to act as a signal sequence and,
15 therefore, within the definition of pre- region of
neurturin is included fragments thereof, usually N-
terminal fragments, that retain the property of being
able to act as a signal sequence, that is to facilitate
co-translational insertion into the membranes of one or
20 more cellular organelles such as endoplasmic reticulum,
mitochondria, golgi, plasma membrane and the like.

       The neurturin signal sequence is followed by a pro-
domain which contains an RXXR proteolytic processing site
immediately before the N-terminal amino acid sequence for
25 the mature neurturin. (human pro- region sequence, SEQ ID
NO:19, Figure 7, amino acids 20 through 95 encoded by the
nucleic acid sequence SEQ ID NO:20, Figure 7 nucleic
acids 58 through 285; mouse pro- region sequence, SEQ ID
NO:22, Figure 8, amino acids 19 through 95 encoded by
30 nucleic acid sequence SEQ ID NO:21, Figure 8, nucleic
acids 58 through 285).

       The neurturin pre- and pro- regions together
comprise a pre-pro sequence identified as the human pre-
pro sequence (SEQ ID NO:23, Figure 7, amino acids 1
35 through 95 encoded by SEQ ID NO:25, nucleic acids 1
through 285) and the mouse pre-pro sequence (SEQ ID

NO:24, Figure 8, amino acids 1 through 95 encoded by SEQ
ID NO:26, nucleic acids 1 through 285). The pre- region
sequences and pro- region sequences as well as the pre-
pro region sequences can be identified and obtained for
5 non-human mammalian species and for non-mammalian species
by virtue of the sequences being contained within the
pre-pro neurturin as defined herein. It is believed that
persephin is similarly associated with pre- and pro-
regions to constitute a pre-pro persephin sequence.

10   Using the above landmarks, the mature, secreted
neurturin molecule is predicted to be approximately 11.5
kD which is likely to form a disulfide linked homodimer
of approximately 23 kD by analogy to other members of the
TGF-ß family. The predicted approximately 23 kD protein
15 is consistent with the 25 kD neurturin protein purified
from CHO cell conditioned media being a homodimer. The
inventors herein have detected an approximately 11.5 kD
neurturin protein from conditioned medium of Chinese
hamster ovary cells transfected with the neurturin
20 expression vector (pCMV-NTN-3-1) using SDS-PAGE under
reducing conditions and this protein is thought to be the
monomer.

      As discussed above, a mature persephin molecule
predicted on the basis of homology to neurturin would
25 contain 5 amino acids upstream from the first canonical
framework cysteine thus having 96 amino acids and a
predicted molecular mass of 10.4 kD. A mature persephin
molecule based upon homology to GDNF would contain 43
amino acids upstream from the first canonical framework
30 cysteine thus having 134 amino acids and a predicted
molecular mass of 14.5 kD.

      The nucleotide sequences of neurturin pre- and/or
pro- regions or similar regions that are believed to be
associated with persephin DNA can be used to construct
35 chimeric genes with the coding sequences of other growth
factors or proteins and, similarly, chimeric genes can be

constructed from the coding sequence of neurturin coupled
to sequences encoding pre- and/or pro- regions from genes
for other growth factors or proteins. (Booth et al.,
Gene 146:303-8, 1994; Ibanez, *Gene 146*:303-8, 1994;

5 Storici et al., *FEBS Letters 337*:303-7, 1994; Sha et al *J
Cell Biol. 114*:827-839, 1991 which are incorporated by
reference). Such chimeric proteins can exhibit altered
production or expression of the active protein species.

A preferred neurturin has been identified and

10 isolated in purified form from medium conditioned by CHO
cells. Also preferred is neurturin prepared by
recombinant DNA technology. Similarly, a preferred
persephin according to the present invention is prepared
by recombinant DNA technology.

15   By "pure form" or "purified form" or "substantially
purified form" it is meant that a persephin or neurturin
composition is substantially free of other proteins which
are not persephin or neurturin, respectively.

Recombinant persephin or neurturin may be made by

20 expressing the DNA sequences encoding persephin or
neurturin, respectively, in a suitable transformed host
cell. Using methods well known in the art, the DNA
encoding persephin or neurturin may be linked to an
expression vector, transformed into a host cell and

25 conditions established that are suitable for expression
of persephin or neurturin, respectively, by the
transformed cell.

Any suitable expression vector may be employed to
produce recombinant human persephin or recombinant human

30 neurturin such as, for example, the mammalian expression
vector pCB6 (Brewer, *Meth Cell Biol 43*:233-245, 1994) or
the *E. coli* pET expression vectors, specifically, pET-30a
(Studier et al., *Methods Enzymol 185*:60-89, 1990 which is
incorporated by reference) both of which were used

35 herein. Other suitable expression vectors for expression
in mammalian and bacterial cells are known in the art as

are expression vectors for use in yeast or insect cells.
Baculovirus expression systems can also be employed.

Persephin or neurturin may be expressed in the
monomeric units or such monomeric form may be produced by
5   preparation under reducing conditions. In such instances
refolding and renaturation can be accomplished using one
of the agents noted above that is known to promote
dissociation/association of proteins. For example, the
monomeric form can be incubated with dithiothreitol
10 followed by incubation with oxidized glutathione disodium
salt followed by incubation with a buffer containing a
refolding agent such as urea.

In the case of neurturin, by analogy with the N-
terminal sequence and internal fragments of the neurturin
15 purified from CHO cell conditioned medium, the mature
mouse sequence was deduced and from this the mature human
form was predicted using the sequence from the human
gene. The amino acid sequence of the mature human form
is as shown in Figure 5 (hNTN, SEQ ID NO:1). The
20 material purified from CHO cell conditioned medium is
considered to be mature neurturin and may exist as a
dimer or other multimer and may be glycosylated or
chemically modified in other ways.

Persephin, like neurturin, may also exist as a dimer
25 or other multimer and may be glycosylated or chemically
modified in other ways.

As noted above, the mouse and human nucleic acid
sequences suggest that neurturin is initially translated
as a pre-pro polypeptide and that proteolytic processing
30 of the signal sequence and the "pro" portion of this
molecule results in the mature sequence, referenced
herein as "mature neurturin", as obtained from medium
condition by CHO cells and as exists in human and in non-
human species in homologous form. Neurturin, therefore,
35 includes any and all "mature neurturin" sequences from
human and non-human species and any and all pre-pro

36

neurturin polypeptides that may be translated from the
neurturin gene.

As with neurturin, the persephin in the present
invention also includes any and all mature persephin
5   sequences from human and non-human species and any and
all pre-pro persephin polypeptides that may be translated
from the persephin gene.

It is believed that the coding sequence for the pre-
pro-neurturin polypeptide begins at the first ATG codon
10 encoding methionine at the 5' end of the clone (position
1 in figure 9) which is positioned in the same reading
frame as the sequence encoding the amino acid sequences
obtained from the purified neurturin. Downstream from
the first codon is the largest open reading frame
15 containing the coding sequence for the pre- and pro-
regions followed by the coding sequence for the mature
mouse neurturin.

Sequence analysis of the murine neurturin genomic
clones identified a 0.5 kb intron located between
20 nucleotide 169 and 170 of the pre-pro neurturin from the
cDNA clones. This intron is located in the coding
sequence of the pro- region of the pre-pro-neurturin
protein. Thus, it is believed that the mouse neurturin
gene contains at least two exons, one of which contains
25 the coding sequences upstream from the splice site and
the other contains the coding sequence downstream (Figure
8, SEQ ID NO:29, SEQ ID NO:30). It is known that the
gene for GDNF contains an intron located at an analogous
position and an alternately spliced form of GDNF has been
30 detected by RT-PCR experiments (Suter-Crazzolara and
Unsicker, *Neuroreport* 5: 2486-2488, 1994 which is
incorporated by reference). This alternate form results
from the use of a splice site in the second coding exon
located 78 bp 3' to the original splice site reported.
35 The alternately spliced form encodes a GDNF protein with
a deletion of 26 amino acids relative to the originally

reported form. The two forms are expressed in different
ratios in different tissues. We have not detected
alternately spliced forms of neurturin in RT-PCR and RACE
experiments using mouse Pl brain and Pl liver cDNAs. The

5   possibility exists, however, that alternate splice sites
in the neurturin gene may be utilized in different
tissues.

       The coding sequence of the human neurturin cDNA has
been deduced from the sequence of the human neurturin

10 genomic clones. The coding sequence of the human cDNA,
like that of the mouse cDNA, is interrupted by an intron
between nucleotides 169 and 170 of the coding sequence.
Thus, the human neurturin gene is believed to contain at
least two exons, one of which contains the coding

15 sequence upstream from the splice site and the other
contains the coding sequence downstream (Figure 7, SEQ ID
NO:27, SEQ ID NO:28). The splice sites at the
intron-exon junctions of the human and mouse genes have
been conserved.

20   From the deduced amino acid sequence of human
neurturin, the earlier predicted N-terminal sequence lies
between positions 286 and 339 and the predicted internal
sequences lie between positions 385 and 417, positions
474 and 533, and positions 547 and 576. The TGA stop

25 codon at positions 592-594 terminate the open reading
frame.

       The predicted length of the purified pre-pro
neurturin is 197 amino acid residues for the human pre-
pro neurturin (SEQ ID NO:7) and 195 amino acid residues

30 for the mouse pre-pro neurturin (SEQ ID NO:8). The
predicted molecular weight of this polypeptide is 22.2 kD
for mouse and 22.4 kd for human. The predicted length of
the purified neurturin is 100 amino acid residues and its
predicted monomeric molecular weight is 11.5 kD. There

35 are no N-linked glycosylation sites, however, potential
O-linked glycosolation sites occur at amino acid residues

38

in positions 18, 26, 80, 86 and 95 in human neurturin.
Glycosylation at any one or combination of these sites
would increase the molecular weight of the molecule.

In the case of persephin, there are no N-linked
5   glycosylation sites in the region between the first and
seventh canonical framework cysteines (SEQ ID NO:79) nor
are there any N-linked glyosylation sites in a mature
persephin molecule predicted on the basis of homology to
neurturin (SEQ ID NO:80). In a mature persephin molecule
10   based upon homology to GDNF there are two potential N-
linked glycosylation sites in the 43 amino acids upstream
from the first canonical framework cysteine at positions
31 and 32 in SEQ ID NO:81 (corresponding to positions 39
and 40 in the sequence as shown in Figure 11).

15   Potential O-linked glycosylation sites occur in
persephin in the region between the first and seventh
canonical framework cysteines at positions 5, 7, 19, 31,
38, 41, 62, 63, 68 and 83 in SEQ ID NO:79 (Figure 12) and
in a mature persephin molecule predicted on the basis of
20   homology to neurturin (SEQ ID NO:80) there is one
additional potential O-linked glycosylation site one
residue upstream from the first canonical framework
cysteine (position 51 in the sequence as shown in Figure
11). In a mature persephin molecule based upon homology
25   to GDNF there are five potential O-linked glycosylation
sites in the 43 amino acids upstream from the first
canonical framework cysteine at positions 9, 15, 18, 22
and 43 in SEQ ID NO:81 (corresponding to positions 17,
23, 26, 30 and 51 in the sequence as shown in Figure 11)
30   along with the ten potential O-linked glycosylation sites
noted above in the region between the canonical framework
cysteines (corresponding to positions 48, 50, 62, 74, 81,
84, 105, 106, 111 and 126 in SEQ ID NO:81 and positions
56, 58, 70, 82, 89, 92, 113, 114, 119 and 134 in the
35   sequence as shown in Figure 11).

39

        Different possible cleavage sites may be present in
the pre-pro-neurturin sequence.   The amino acid sequence
of the mature mouse neurturin (Figure 5, SEQ ID NO:2) is
predicted from alignment with the N-terminal amino acid
5   sequence of the purified chinese hamster neurturin.   A
four residue RRAR cleavage site (amino acids 92-95) is
found immediately before the predicted N-terminal amino
acid of mature mouse neurturin.   This RRAR sequence fits
the RXXR consensus sequence at which members of the TGF-ß
10 superfamily are usually cleaved.   This putative RRAR
cleavage sequence is conserved in human neurturin.
However, the mature human neurturin is predicted to have
a two amino acid N-terminal extension relative to mature
mouse neurturin when cleaved at this sequence.   Since
15 neurturin contains other sequences which fit the RXXR
consensus (for example the sequence RRRR at amino acids
90-93) and the specificities of proteases involved in
this cleavage are not completely understood, the
possibility exists that in some situations, neurturin is
20 cleaved at sites other than the above RRAR sequence, and
the mature neurturin protein may have a variable number
of amino acids preceding the cysteine residue at position
101 in the mouse sequence (pre-pro protein) and position
103 in the human sequence.   Such alternate cleavage sites
25 could be utilized differently among different organisms
and among different tissues of the same organism.   The
N-terminal amino acids preceding the first of the seven
conserved cysteines in the mature forms of members of the
TGF-ß family vary greatly in both length and sequence.
30 Furthermore, insertion of a ten amino acid sequence two
residues upstream of the first conserved cysteine does
not affect the known biological activities of one family
member, dorsalin (Basler, K., Edlund, T., Jessell,T.M.,
and Yamada, T., (1993) Cell 73:687-702).   Thus neurturin
35 proteins which contain sequences of different lengths
preceding the cysteine 101 in mouse and cysteine 103 in

40

human would be likely to retain their biological
activity.

It is also believed that persephin proteins which
contain sequences of different lengths preceding the
.5   first cysteine (residue No. 1 of mouse persephin in
Figure 12 and residue No. 1 of rat persephin in Figure
14) would be likely to retain their biological activity.

The inventors herein believe that at a minimum the
sequence of neurturin that will show biological activity
10   will contain the sequence beginning at cysteine 103 and
ending at cysteine 196 for human neurturin (Figure 7, SEQ
ID NO:31) and beginning at cysteine 101 and ending at
cysteine 194 for mouse neurturin (Figure 7, SEQ ID
NO:32). Thus, within the scope of the neurturin
15   polypeptides are amino acid sequences containing SEQ ID
NO:31 and amino acid sequences containing SEQ ID NO:32
and nucleic acid sequences encoding these amino acid
sequences.

Similarly, the inventors herein believe that, at a
20   minimum, the sequence of persephin that will show
biological activity will contain the sequence beginning
at cysteine 1 and ending at cysteine 87 for mouse
persephin (Figure 12, SEQ ID NO:79) and beginning at
cysteine 1 and ending at cysteine 87 for rat persephin
25   (Figure 14, SEQ ID NO:82). Thus, within the scope of
persephin of the present invention are amino acid
sequences containing SEQ ID NO:79 and amino acid
sequences containing SEQ ID NO:82 and nucleic acid
sequences encoding these amino acid sequences.

30   The present invention also encompasses nucleic acid
sequences including sequences that encode mouse and rat
persephin (Figures 11 and 14) as well as human persephin
in the same manner that neurturin includes human and
mouse neurturin nucleic acid sequences (Figures 7 and 8).
35   Also included within the scope of this invention are
sequences that are substantially the same as the nucleic

41

acid sequences encoding persephin or neurturin,
respectively. Such substantially the same sequences may,
for example, be substituted with codons more readily
expressed in a given host cell such as *E. coli* according

5   to well known and standard procedures. Such modified
nucleic acid sequences are included within the scope of
this invention.

Specific nucleic acid sequences can be modified by
those skilled in the art and, thus, all nucleic acid

10 sequences which encode for the amino acid sequences of
pre-pro neurturin or persephin or the pre- region or the
pro- region of neurturin or persephin can likewise be so
modified. The present invention thus also includes
nucleic acid sequence which will hybridize with all such

15 nucleic acid sequences -- or complements of the nucleic
acid sequences where appropriate -- and encode for a
polypeptide having cell survival or growth promoting
activity. The present invention also includes nucleic
acid sequences which encode for polypeptides that have

20 survival or growth promoting activity and that are
recognized by antibodies that bind to neurturin or by
antibodies that bind to persephin.

The present invention also encompasses vectors
comprising expression regulatory elements operably linked

25 to any of the nucleic acid sequences included within the
scope of the invention. This invention also includes
host cells -- of any variety -- that have been
transformed with vectors comprising expression regulatory
elements operably linked to any of the nucleic acid

30 sequences included within the scope of the present
invention.

Methods are also provided herein for producing
neurturin or persephin. Preparation can be by isolation
from conditioned medium from a variety of cell types so

35 long as the cell type produces neurturin or persephin. A
second and preferred method involves utilization of

42

recombinant methods by isolating a nucleic acid sequence
encoding neurturin or persephin, cloning the sequence
along with appropriate regulatory sequences into suitable
vectors and cell types, and expressing the sequence to
5   produce neurturin or persephin.

A mammalian gene family comprised of four
neurotrophic factors has been identified including nerve
growth factor (NGF), brain derived neurotrophic factor
(BDGF), neurotrophin-3 (NT-3), and neurotrophin-4/5 (NT-
10 4/5).   These factors share approximately 60 percent
nucleic acid sequence homology (Tuszynski and Gage, *Ann
Neurol* 35:S9-S12, 1994 which is incorporated by
reference). The persephin protein and the neurturin
protein display no significant homology to the NGF family
15 of neurotrophic factors. Either persephin or neurturin
shares less than about 20% homology with the TGF-ß
superfamily of growth factors. However, both persephin
and neurturin show approximately 40% sequence identity
with GDNF and approximately 50% sequence identity with
20 each other. In particular, the positions of the seven
cysteine residues present in persephin, neurturin and
GDNF are nearly exactly conserved. The inventors herein
believe that other unidentified genes may exist that
encode proteins that have substantial amino acid sequence
25 homology to persephin, neurturin and GDNF and which
function as growth factors selective for the same or
different tissues and the same or different biological
activities and may act at the same or different
receptors. A different spectrum of activity with respect
30 to tissues affected and/or response elicited could result
from preferential activation of different receptors by
different family members as is known to occur with
members of the NGF family of neurotrophic factors
(Tuszynski and Gage, 1994, *supra*).

35   As a consequence of members of a particular gene
family showing substantial conservation of amino acid

43

sequence among the protein products of the family
members, there is considerable conservation of sequences
at the DNA level. This forms the basis for a new
approach for identifying other members of the gene family
5   to which GDNF, neurturin and persephin belong. The
method used for such identification is cross-
hybridization using nucleic acid probes derived from one
family member to form a stable hybrid duplex molecule
with nucleic acid sequence from different members of the
10 gene family or to amplify nucleic acid sequences from
different family members. (see for example, Kaisho et al.
FEBS Letters 266:187-191, 1990 which is incorporated by
reference). The sequence from the different family
member may not be identical to the probe, but will,
15 nevertheless be sufficiently related to the probe
sequence to hybridize with the probe. Alternatively, PCR
using primers from one family member can be used to
identify additional family members.

The above approaches have not heretofore been
20 successful in identifying other gene family members
because only one family member, GDNF was known. With the
identification of neurturin in copending application
Serial No. 08/519,777, however, unique new probes and
primers can be made that contain sequences from the
25 conserved regions of this gene family. The same
conserved regions are also found in the third family
member, persephin. In particular, three conserved
regions have been identified herein which can be used as
a basis for constructing new probes and primers. The new
30 probes and primers made available from the work with
neurturin and persephin make possible this powerful new
approach which can now successfully identify other gene
family members. Using this new approach, one may screen
for genes related to GDNF, neurturin and persephin in
35 sequence homology by preparing DNA or RNA probes based
upon the conserved regions in the GDNF and neurturin

44

molecules.   Therefore, one embodiment of the present
invention comprises probes and primers that are unique to
or derived from a nucleotide sequence encoding such
conserved regions and a method for identifying further
5   members of the neurturin-persephin-GDNF gene family.
        Conserved-region amino acid sequences have been
identified herein to include Val-Xaa$_1$-Xaa$_2$-Leu-Gly-Leu-
Gly-Tyr where Xaa$_1$ is Ser, Thr or Ala and Xaa$_2$ is Glu or
Asp (SEQ ID NO:108); Glu-Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Arg-Tyr-Cys-
10 Xaa$_4$-Gly-Xaa$_5$-Cys in which Xaa$_1$ is Thr, Glu or lys, Xaa$_2$ is
Val, Leu or Ile, Xaa$_3$ is Leu or Ile, Xaa$_4$ is Ala or Ser,
and Xaa$_5$ is Ala or Ser, (SEQ ID NO:113); and Cys-Cys-Xaa$_1$-
Pro-Xaa$_2$-Xaa$_3$-Xaa$_4$-Xaa$_5$-Asp-Xaa$_6$-Xaa$_7$-Xaa$_8$-Phe-Leu-Asp-Xaa$_9$
in which Xaa$_1$ is Arg or Gln, Xaa$_2$ is Thr or Val or Ile,
15 Xaa$_3$ is Ala or Ser, Xaa$_4$ is Tyr or Phe, Xaa$_5$ is Glu, Asp
or Ala, Xaa$_6$ is Glu, Asp or no amino acid, Xaa$_7$ is val or
leu, Xaa$_8$ is Ser or Thr, and Xaa$_9$ is Asp or Val (SEQ ID
NO:114). Nucleotide sequences containing a coding
sequence for the above conserved sequences or fragments
20 of the above conserved sequences can be used as probes.
Exemplary probe and primer sequences encoding amino acid
sequences and SEQ ID NOS:125-129; primers whose reverse
complementary sequences encode amino acid sequences SEQ
ID NO:126, SEQ ID NO:127, SEQ ID NO:130; and, in
25 particular, nucleotide sequences, SEQ ID NOS:115-124.
Additional primers based upon GDNF and neurturin include
nucleic acid sequences encoding amino acid sequences, SEQ
ID NO:33, SEQ ID NO:36, SEQ ID NO:40 and SEQ ID NO:41;
primers whose reverse complementary sequences encode SEQ
30 ID NO:37, SEQ ID NO:38 and SEQ ID NO 39; and, in
particular, nucleic acid sequences, SEQ ID NOS:42-48.
        Hybridization using the new probes from conserved
regions of the nucleic acid sequences would be performed
under reduced stringency conditions. Factors involved in
35 determining stringency conditions are well known in the
art (for example, see Sambrook et al., *Molecular Cloning,*

2nd Ed., 1989 which is incorporated by reference).
Sources of nucleic acid for screening would include
genomic DNA libraries from mammalian species or cDNA
libraries constructed using RNA obtained from mammalian

5   cells cloned into any suitable vector.

PCR primers would be utilized under PCR conditions
of reduced annealing temperature which would allow
amplification of sequences from gene family members other
than GDNF, neurturin and persephin. Sources of nucleic

10 acid for screening would include genomic DNA libraries
from mammalian species cloned into any suitable vector,
cDNA transcribed from RNA obtained from mammalian cells,
and genomic DNA from mammalian species.

DNA sequences identified on the basis of

15 hybridization or PCR assays would be sequenced and
compared to GDNF, neurturin and persephin. The DNA
sequences encoding the entire sequence of the novel
factor would then be obtained in the same manner as
described herein. Genomic DNA or libraries of genomic

20 clones can also be used as templates because the
intron/exon structures of GDNF and neurturin are
conserved and coding sequences of the mature proteins are
not interrupted by introns.

Using this approach as described above, the primers

25 designed from the conserved regions of neurturin and GDNF
have been used to identify and obtain the sequence of the
new family member described herein, persephin.
Degenerate primers designed from persephin, neurturin and
GDNF can be further used to identify and obtain

30 additional family members.

It is believed that all GDNF-neurturin-persephin
family members will have a high degree of sequence
identity with one or more of the three identified family-
member consensus regions in the portion of the sequence

35 between the first and seventh canonical framework
cysteines (see Figure 12). In particular, a new family

46

member is anticipated to have at least a 62.5% identity
with the consensus region octapeptide, Val-Xaa$_1$-Xaa$_2$-Leu-
Gly-Leu-Gly-Tyr where Xaa$_1$ is Ser, Thr or Ala and Xaa$_2$ is
Glu or Asp (SEQ ID NO:108) or at least a 62.5 percent

5   sequence identity with the consensus region octapeptide,
Phe-Arg-Tyr-Cys-Xaa$_1$-Gly-Xaa$_2$-Cys where Xaa$_1$ and Xaa$_2$ are
alanine or serine (SEQ ID NO:109) or at least a 50
percent sequence identity with the consensus region
octapeptide, Asp-Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Leu-Asp-Xaa$_4$ where Xaa$_1$

10   is aspartic acid or glutamic acid or no amino acid, Xaa$_2$
is valine or leucine, Xaa$_3$ is serine or threonine; and
Xaa$_4$ is valine or aspartic acid (SEQ ID NO:110). The
inventors herein believed that any new family member will
have 28 amino acids in the aligned sequence between the

15   first and seventh canonical framework cysteine residues
as set forth in Figure 15 with residues numbered from the
N-terminal end of the family member aligned sequence
being (1) Cys, (3) Leu, (10) Val, (13) Leu, (14) Gly,
(15) Leu, (16) Gly, (17) Tyr, (21) Glu, (25) Phe, (26)

20   Arg, (27) Tyr, (28) Cys, (30) Gly, (32) Cys, (44) Leu,
(47) Leu, (58) Cys, (59) Cys (61) Pro, (66) Asp, (69)
Phe, (70) Leu, (71) Asp, (83) Ser, (84) Ala, (87) Cys,
and (89) Cys, however, it is possible that there may be
as many as three mismatches.

25   Although neurturin has been purified on the basis of
its ability to promote the survival of a particular
neuronal type, this factor will act on other neuronal
cell types as well. For example, neurturin is shown
herein to promote the survival of nodose sensory ganglia

30   neurons (see Example 3). Neurturin is also likely to
promote the survival of non-neuronal cells. Indeed, all
the growth factors isolated to date have been shown to
act on many different cell types (for example see Scully
and Otten, *Cell Biol Int* 19:459-469, 1005; Hefti,

35   *Neurotrophic Factor Therapy* 25:1418-1435, 1994 which are
incorporated by reference). It is known that NGF acts on

sympathetic neurons, several types of sensory neurons and
certain populations of CNS neurons. GDNF, which is more
closely related to neurturin, has been shown to act on
dopaminergic, sympathetic, motor and several sensory
5 neurons (Henderson et al. *supra*, 1994; Miles et al, *J
Cell Biol* 130:137-148, 1995; Yan et al, *Nature* 373:341-
344, 1995; Lin et al, *Science* 260:1130-1132, 1993; Trupp
et al, *J Cell Biol* 130:137-148, 1995; Martin et al *Brain
Res* 683:172-178, 1995; Bowenkamp st al *J Comp Neurol*
10 355:479-489, 1995 which are incorporated by reference).
Thus, it is likely that in addition to peripheral
sympathetic and sensory neurons, neurturin can act on a
wide variety of central and peripheral neuronal cell
types.

15   On the basis of the structural similarities of
persephin to the sequences of neurturin and GDNF,
persephin is also believed to promote the survival and
growth of neuronal as well as non-neuronal cells.
Indeed, as noted above, all the growth factors isolated
20 to date have been shown to act on many different cell
types (Scully and Otten, *Cell Biol Int* 19:459-469, 1005;
Hefti, *Neurotrophic Factor Therapy* 25:1418-1435, 1994).
Furthermore, the inventors herein have identified brain
and heart tissues as tissues expressing persephin, which
25 further supports the conclusion that persephin can act to
promote survival and growth in a variety of neuronal and
non-neuronal cells.

As an example of the actions of neurotrophic factors
on non-neuronal tissues, the prototypical neurotrophic
30 factor, NGF, also acts upon mast cells to increase their
number when injected into newborn rats (Aloe, *J
Neuroimmunol* 18:1-12, 1988). In addition, mast cells
express the trk receptor and respond to NGF such that NGF
is a mast cell secretogogue and survival promoting factor
35 (Horigome et al., *J Biol Chem* 269:2695-2707, 1994 which
is incorporated by reference). Moreover, members of the

48

TGF-β superfamily act on many cell types of different
function and embryologic origin.

The inventors herein have identified several non-
neuronal tissues in which neurturin is expressed
5   including blood, bone marrow, neonatal liver and mast
cells. This suggests a role for neurturin in
hematopoiesis, inflammation, allergy, and cardiomyopathy.

Similarly, the inventors herein have identified
brain and heart as tissues in which persephin is
10 expressed and it is further believed that persephin is
expressed in a number of other neuronal and non-neuronal
tissues. Thus, persephin may also have a role in
hematopoiesis, inflammation, allergy and
cardiomyopathies.

15 Neurotrophic factors of the NGF family are thought
to act through factor-specific high affinity receptors
(Tuszynski and Gage, 1994, *supra*). Only particular
portions of the protein acting at a receptor site are
required for binding to the receptor. Such particular
20 portions or discrete fragments can serve as an agonist
where the substance activates the receptor to elicit the
promoting action on cell survival and growth and
antagonists to neurturin or persephin where they bind to,
but do not activate, the receptor or promote survival and
25 growth. Such portions or fragments that are agonists and
those that are antagonists are also within the scope of
the present invention.

Synthetic, pan-growth factors can also be
constructed by combining the active domains of persephin
30 or neurturin with the active domains of one or more other
growth factors. (For example, see Ilag et al., *Proc Nat'l
Acad Sci* 92:607-611, 1995 which is incorporated by
reference). These pan-growth factors would be expected
to have the combined activities of neurturin or persephin
35 and the one or more other growth factors. As such they
are believed to be potent and multispecific growth

49

factors that are useful in the treatment of a wide
spectrum of degenerative diseases and conditions
including conditions that can be treated by any and all
of the parent factors from which the active domains were
5 obtained. Such pan-growth factors might also provide
synergistic effects beyond the activities of the parent
factors (Barres et al., supra).

     Pan-growth factors within the scope of the present
invention can also include chimeric or hybrid
10 polypeptides that are constructed from portions of
fragments of at least two growth factors. Growth factors
of the TGF-ß superfamily are structurally related having
highly conserved sequence landmarks whereby family
members are identified. In particular, seven canonical
15 framework cysteine residues are nearly invariant in
members of the superfamily (Kingsley, Genes & Dev 8:133-
146, 1994 which is incorporated by reference)(see Figure
17). Chimeric polypeptide molecules can, therefore, be
constructed from a sequence that is substantially
20 identical to a portion of either the persephin or the
neurturin molecule, up to one or more crossover points,
and one or more sequences each of which is substantially
identical with a portion of another TGF-ß superfamily
member extending on the other side of the corresponding
25 one or more crossover points. For example, a portion of
the amino terminal end of the persephin polypeptide can
be combined with a portion of the carboxy terminal end of
a neurturin polypeptide or alternatively a portion of the
amino terminal end of a neurturin polypeptide can be
30 combined with a portion of the carboxy terminal end of a
persephin polypeptide. Such portions of neurturin or
persephin polypeptides are preferably from about 5 to
about 95, more preferably from about 10 to about 90,
still more preferably from about 20 to about 80 and most
35 preferably from about 30 to about 70 contiguous amino
acids and such portions of another, non-persephin or, as

the case may be, non-neurturin TGF-ß superfamily member
are preferably from about 5 to about 95, more preferably
from about 10 to about 90, still more preferably from
about 20 to about 80 and most preferably from about 30 to

5   about 70 contiguous amino acids. For example, a
particular crossover point might be between the third and
fourth canonical framework cysteine residues. One such
exemplary construct would contain at the 5' end a
sequence comprised of a persephin sequence from residue 1

10 through the third canonical framework cysteine residue 37
and up to a cross-over point somewhere between residue 37
and residue 63 but not including the fourth canonical
framework cysteine residue 64 (for reference, see mature
persephin, SEQ ID NO:80). The 3' end of the hybrid

15 construct would constitute a sequence derived from
another TGF-ß superfamily member such as, for example,
neurturin which is another TGF-ß superfamily member that
is closely related to persephin. Using neurturin as the
other TGF-ß family member, the hybrid construct beyond

20 the crossover point would be comprised of a sequence
beginning at the desired crossover point in the neurturin
sequence between the third canonical framework cysteine
residue 37 and the fourth canonical framework cysteine
residue 67 of neurturin and continuing through residue

25 100 at the 3' end of neurturin (for alignment, see figure
12). A second exemplary hybrid construct would be
comprised of residue 1 through a crossover point between
residues 37 and 67 of neurturin contiguously linked with
residues from the crossover point between residues 37 and

30 64 through residue 96 of persephin. The above constructs
with persephin and neurturin are intended as examples
only with the particular TGF-ß family member being
selected from family members including but not limited to
transforming growth factor-ß1 (TGFß1), transforming

35 growth factor-ß2 (TGFß2), transforming growth factor-ß3
(TGFß3), inhibin ß A (INHßA), inhibin ß B (INHßB), the

51

nodal gene (NODAL), bone morphogenetic proteins 2 and 4
(BMP2 and BMP4), the *Drosophila decapentaplegic* gene
(*dpp*), bone morphogenetic proteins 5-8 (BMP5, BMP6, BMP7
and BMP8), the *Drosophila* 60A gene family (60A), bone

5   morphogenetic protein 3 (BMP3), the *Vgl* gene, growth
differentiation factors 1 and 3 (GDF1 and GDF3), dorsalin
(drsln), inhibin α (INHα), the *MIS* gene (MIS), growth
factor 9 (GDF-9), glial-derived neurotropic growth factor
(GDNF), neurturin (NTN) and persephin (see Figure 16).

10 In addition, the crossover point can be any residue
between the first and seventh canonical framework
cysteines molecules of neurturin and the particular other
family member. Furthermore, additional crossover points
can be used to incorporate any desired number of

15 persephin portions or fragments with portions or
fragments of any one or more other family members.

In constructing a particular chimeric molecule, the
portions of persephin and portions of the other, non-
persephin growth factor are amplified using PCR, mixed

20 and used as template for a PCR reaction using the forward
primer from one and the reverse primer from the other of
the two component portions of the chimeric molecule.
Thus, for example a forward and reverse primers are
selected to amplify the portion of persephin from the

25 beginning to the selected crossover point between the
third and fourth canonical cysteine residues using a
persephin plasmid as template. A forward primer with a
5' portion overlapping with the persephin sequence and a
reverse primer are then used to amplify the portion of

30 the other, non-persephin growth factor member of the TGF-
β superfamily from the corresponding crossover point
through the 3' end using a plasmid template containing
the coding sequence for the non-persephin TGF-β family
member. The products of the two PCR reactions are gel

35 purified and mixed together and a PCR reaction performed.
Using an aliquot of this reaction as template a PCR

52

reaction is performed using the persephin forward primer and the reverse primer for the non-persephin growth factor. The product is then cloned into an expression vector for production of the chimeric molecule.

5   Chimeric growth factors would be expected to be effective in promoting the growth and development of cells and for use in preventing the atrophy, degeneration or death of cells, particular in neurons. The chimeric polypeptides may also act as a receptor antagonists of

10 one or both of the full length growth factors from which the chimeric polypeptide was constructed or as an antagonist of any other growth factor that acts at the same receptor or receptors.

The present invention also includes therapeutic or

15 pharmaceutical compositions comprising persephin or neurturin in an effective amount for treating patients with cellular degeneration or dysfunction and a method comprising administering a therapeutically effective amount of neurturin or persephin. These compositions and

20 methods are useful for treating a number of degenerative diseases. Where the cellular degeneration involves neuronal degeneration, the diseases include, but are not limited to peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease,

25 Huntington's disease, ischemic stroke, acute brain injury, acute spinal chord injury, nervous system tumors, multiple sclerosis, peripheral nerve trauma or injury, exposure to neurotoxins, metabolic diseases such as diabetes or renal dysfunctions and damage caused by

30 infectious agents. In particular, the ability of persephin to promote survival in mesencephalic cells suggests an applicability of this growth factor in treating neuronal degenerative diseases of the CNS such as Parkinson's disease.

35   Where the cellular degeneration involves bone marrow cell degeneration, the diseases include, but are not

53

limited to disorders of insufficient blood cells such as, for example, leukopenias including eosinopenia and/or basopenia, lymphopenia, monocytopenia, neutropenia, anemias, thrombocytopenia as well as an insufficiency of

5   stem cells for any of the above. The cellular degeneration can also involve myocardial muscle cells in diseases such as cardiomyopathy and congestive heart failure. The above cells and tissues can also be treated for depressed function.

10   The compositions and methods herein can also be useful to prevent degeneration and/or promote survival in other non-neuronal tissues as well. One skilled in the art can readily determine using a variety of assays known in the art for identifying whether neurturin or persephin

15 would be useful in promoting survival or functioning in a particular cell type.

        In certain circumstances, it may be desirable to modulate or decrease the amount of persephin or neurturin expressed. Thus, in another aspect of the present

20 invention, persephin or neurturin anti-sense oligonucleotides can be made and a method utilized for diminishing the level of expression of persephin or neurturin, respectively, by a cell comprising administering one or more persephin or neurturin anti-

25 sense oligonucleotides. By persephin or neurturin anti-sense oligonucleotides reference is made to oligonucleotides that have a nucleotide sequence that interacts through base pairing with a specific complementary nucleic acid sequence involved in the

30 expression of persephin or neurturin, respectively, such that the expression of persephin or neurturin is reduced. Preferably, the specific nucleic acid sequence involved in the expression of persephin or neurturin is a genomic DNA molecule or mRNA molecule that contains sequences of

35 the persephin or neurturin gene. This genomic DNA molecule can comprise flanking regions of the persephin

54

or neurturin gene, untranslated regions of persephin or neurturin mRNA, the pre- or pro- portions of the persephin or neurturin gene or the coding sequence for mature persephin or neurturin protein. The term

5 complementary to a nucleotide sequence in the context of persephin or neurturin antisense oligonucleotides and methods therefor means sufficiently complementary to such a sequence as to allow hybridization to that sequence in a cell, i.e., under physiological conditions. The

10 persephin or neurturin antisense oligonucleotides preferably comprise a sequence containing from about 8 to about 100 nucleotides and more preferably the persephin or neurturin antisense oligonucleotides comprise from about 15 to about 30 nucleotides. The persephin or

15 neurturin antisense oligonucleotides can also contain a variety of modifications that confer resistance to nucleolytic degradation such as, for example, modified internucleoside linkages (Uhlmann and Peyman, *Chemical Reviews* 90:543-548, 1990; Schneider and Banner,

20 *Tetrahedron Lett* 31:335, 1990 which are incorporated by reference), modified nucleic acid bases and/or sugars and the like.

The therapeutic or pharmaceutical compositions of the present invention can be administered by any suitable

25 route known in the art including for example intravenous, subcutaneous, intramuscular, transdermal, intrathecal or intracerebral. Administration can be either rapid as by injection or over a period of time as by slow infusion or administration of slow release formulation. For treating

30 tissues in the central nervous system, administration can be by injection or infusion into the cerebrospinal fluid (CSF). When it is intended that neurturin or persephin be administered to cells in the central nervous system, administration can be with one or more agents capable of

35 promoting penetration of neurturin or persephin across the blood-brain barrier.

55

Persephin or neurturin can also be linked or
conjugated with agents that provide desirable
pharmaceutical or pharmacodynamic properties.   For
example, persephin or neurturin can be coupled to any
5 substance known in the art to promote penetration or
transport across the blood-brain barrier such as an
antibody to the transferrin receptor, and administered by
intravenous injection. (See for example, Friden et al.,
Science 259:373-377, 1993 which is incorporated by
10 reference).   Furthermore, persephin or neurturin can be
stably linked to a polymer such as polyethylene glycol to
obtain desirable properties of solubility, stability,
half-life and other pharmaceutically advantageous
properties. (See for example Davis et al. *Enzyme Eng*
15 *4*:169-73, 1978; Burnham, *Am J Hosp Pharm* 51:210-218, 1994
which are incorporated by reference).

The compositions are usually employed in the form of
pharmaceutical preparations.   Such preparations are made
in a manner well known in the pharmaceutical art.   One
20 preferred preparation utilizes a vehicle of physiological
saline solution, but it is contemplated that other
pharmaceutically acceptable carriers such as
physiological concentrations of other non-toxic salts,
five percent aqueous glucose solution, sterile water or
25 the like may also be used.   It may also be desirable that
a suitable buffer be present in the composition.   Such
solutions can, if desired, be lyophilized and stored in a
sterile ampoule ready for reconstitution by the addition
of sterile water for ready injection.   The primary
30 solvent can be aqueous or alternatively non-aqueous.
Persephin or neurturin can also be incorporated into a
solid or semi-solid biologically compatible matrix which
can be implanted into tissues requiring treatment.

The carrier can also contain other pharmaceutically-
35 acceptable excipients for modifying or maintaining the
pH, osmolarity, viscosity, clarity, color, sterility,

56

stability, rate of dissolution, or odor of the
formulation. Similarly, the carrier may contain still
other pharmaceutically-acceptable excipients for
modifying or maintaining release or absorption or
5 penetration across the blood-brain barrier. Such
excipients are those substances usually and customarily
employed to formulate dosages for parenteral
administration in either unit dosage or multi-dose form
or for direct infusion into the cerebrospinal fluid by
10 continuous or periodic infusion.

Dose administration can be repeated depending upon
the pharmacokinetic parameters of the dosage formulation
and the route of administration used.

It is also contemplated that certain formulations
15 containing persephin or neurturin are to be administered
orally. Such formulations are preferably encapsulated
and formulated with suitable carriers in solid dosage
forms. Some examples of suitable carriers, excipients,
and diluents include lactose, dextrose, sucrose,
20 sorbitol, mannitol, starches, gum acacia, calcium
phosphate, alginates, calcium silicate, microcrystalline
cellulose, polyvinylpyrrolidone, cellulose, gelatin,
syrup, methyl cellulose, methyl- and
propylhydroxybenzoates, talc, magnesium, stearate, water,
25 mineral oil, and the like. The formulations can
additionally include lubricating agents, wetting agents,
emulsifying and suspending agents, preserving agents,
sweetening agents or flavoring agents. The compositions
may be formulated so as to provide rapid, sustained, or
30 delayed release of the active ingredients after
administration to the patient by employing procedures
well known in the art. The formulations can also contain
substances that diminish proteolytic degradation and
promote absorption such as, for example, surface active
35 agents.

57

The specific dose is calculated according to the
approximate body weight or body surface area of the
patient or the volume of body space to be occupied. The
dose will also be calculated dependent upon the

5 particular route of administration selected. Further
refinement of the calculations necessary to determine the
appropriate dosage for treatment is routinely made by
those of ordinary skill in the art. Such calculations
can be made without undue experimentation by one skilled

10 in the art in light of the activity of neurturin or GDNF.
With neurturin the activity in target cells data is
disclosed herein and in copending application Serial
Number 08/519,777 and in the case of persephin, the
concentration required for activity at the cellular level

15 is believed to be similar to that of neurturin.
Persephin activity on a particular target cell type can
be determined by routine experimentation. Exact dosages
are determined in conjunction with standard dose-response
studies. It will be understood that the amount of the

20 composition actually administered will be determined by a
practitioner, in the light of the relevant circumstances
including the condition or conditions to be treated, the
choice of composition to be administered, the age,
weight, and response of the individual patient, the

25 severity of the patient's symptoms, and the chosen route
of administration.

In one embodiment of this invention, persephin or
neurturin may be therapeutically administered by
implanting into patients vectors or cells capable of

30 producing a biologically-active form of persephin or
neurturin or a precursor of persephin or neurturin, i.e.
a molecule that can be readily converted to a biological-
active form of neurturin by the body. In one approach
cells that secrete persephin or neurturin may be

35 encapsulated into semipermeable membranes for
implantation into a patient. The cells can be cells that

58

normally express persephin or neurturin or a precursor
thereof or the cells can be transformed to express
persephin or neurturin or a precursor thereof. It is
preferred that the cell be of human origin and that the
5   persephin or neurturin be human persephin or neurturin
when the patient is human. However, the formulations and
methods herein can be used for veterinary as well as
human applications and the term "patient" as used herein
is intended to include human and veterinary patients.
10   Cells can be grown *ex vivo* for use in
transplantation or engraftment into patients (Muench et
al., *Leuk & Lymph* 16:1-11, 1994 which is incorporated by
reference). In another embodiment of the present
invention, persephin or neurturin is used to promote the
15   *ex vivo* expansion of a cells for transplantation or
engraftment. Current methods have used bioreactor
culture systems containing factors such as
erythropoietin, colony stimulating factors, stem cell
factor, and interleukins to expand hematopoietic
20   progenitor cells for erythrocytes, monocytes,
neutrophils, and lymphocytes (Verfaillie, *Stem Cells*
12:466-476, 1994 which is incorporated by reference).
These stem cells can be isolated from the marrow of human
donors, from human peripheral blood, or from umbilical
25   cord blood cells. The expanded blood cells are used to
treat patients who lack these cells as a result of
specific disease conditions or as a result of high dose
chemotherapy for treatment of malignancy (George, *Stem
Cells* 12(Suppl 1):249-255, 1994 which is incorporated by
30   reference). In the case of cell transplant after
chemotherapy, autologous transplants can be performed by
removing bone marrow cells before chemotherapy, expanding
the cells *ex vivo* using methods that also function to
purge malignant cells, and transplanting the expanded
35   cells back into the patient following chemotherapy (for
review see Rummel and Van Zant, *J Hematotherapy* 3:213-

59

218, 1994 which is incorporated by reference). Since
persephin or neurturin is believed to be expressed in the
developing animal in blood, bone marrow and liver,
tissues where proliferation and differentiation of
5   progenitor cells occur, it is believed that persephin or
neurturin can function to regulate the proliferation of
hematopoietic stem cells and the differentiation of
mature hematopoietic cells. Thus, the addition of
persephin or neurturin to culture systems used for *ex*
10 *vivo* expansion of cells could stimulate the rate at which
certain populations of cells multiply or differentiate,
and improve the effectiveness of these expansion systems
in generating cells needed for transplant.

        It is also believed that persephin or neurturin can
15 be used for the *ex vivo* expansion of precursor cells in
the nervous system. Transplant or engraftment of cells
is currently being explored as a therapy for diseases in
which certain populations of neurons are lost due to
degeneration such as, for example, in parkinson's disease
20 (Bjorklund, *Curr Opin Neurobiol* 2:683-689, 1992 which is
incorporated by reference). Neuronal precursor cells can
be obtained from animal or human donors or from human
fetal tissue and then expanded in culture using persephin
or neurturin or other growth factors. These cells can
25 then be engrafted into patients where they would function
to replace some of the cells lost due to degeneration.
Because neurotrophins have been shown to be capable of
stimulating the survival and proliferation of neuronal
precursors cells such as, for example, NT-3 stimulation
30 of sympathetic neuroblast cells (Birren et al., *Develop*
119:597-610, 1993 which is incorporated by reference),
persephin or neurturin could also function in similar
ways during the development of the nervous system and
could be useful in the *ex vivo* expansion of neuronal
35 cells.

60

In a number of circumstances it would be desirable to determine the levels of persephin or neurturin in a patient. The identification of persephin or neurturin along with the present report that persephin and

5 neurturin are expressed by a number of tissues provides the basis for the conclusion that the presence of persephin or neurturin serves a normal physiologic function related to cell growth and survival. Indeed, other neurotrophic factors are known to play a role in

10 the function of neuronal and non-neuronal tissues. (For review see Scully and Otten, *Cell Biol Int* *19*:459-469, 1995; Otten and Gadient, *Int J Devl Neurosciences* *13*:147-151, 1995 which are incorporated by reference). Endogenously produced persephin or neurturin may also

15 play a role in certain disease conditions, particularly where there is cellular degeneration such as in neurodegenerative conditions or diseases. Other neurotrophic factors are known to change during disease conditions. For example, in multiple sclerosis, levels

20 of NGF protein in the cerebrospinal fluid are increased during acute phases of the disease (Bracci-Laudiero et al., *Neuroscience Lett* *147*:9-12, 1992 which is incorporated by reference) and in systemic lupus erythematosus there is a correlation between inflammatory

25 episodes and NGF levels in sera (Bracci-Laudiero et al. *NeuroReport* *4*:563-565, 1993 which is incorporated by reference).

Given that neurturin is expressed in blood cells, bone marrow and mast cells, and it is believed that

30 persephin is similarly expressed, it is likely that the level of persephin or neurturin may be altered in a variety of conditions and that quantification of persephin or neurturin levels would provide clinically useful information. Furthermore, in the treatment of

35 degenerative conditions, compositions containing either persephin or neurturin or both can be administered and it

61

would likely be desirable to achieve certain target
levels of persephin and/or neurturin, as the case may be,
in sera, in cerebrospinal fluid or in any desired tissue
compartment. It would, therefore, be advantageous to be
5   able to monitor the levels of the particular growth
factor, persephin or neurturin, in a patient.
Accordingly, the present invention also provides methods
for detecting the presence of persephin or for detecting
the presence of neurturin in a sample from a patient.
10   The term "detection" as used herein in the context
of detecting the presence of persephin or neurturin in a
patient is intended to include the determining of the
amount of persephin or neurturin or the ability to
express an amount of persephin or neurturin in a patient,
15 the distinguishing of persephin or neurturin from other
growth factors, the estimation of prognosis in terms of
probable outcome of a degenerative disease and prospect
for recovery, the monitoring of the persephin or
neurturin levels over a period of time as a measure of
20 status of the condition, and the monitoring of persephin
or neurturin levels for determining a preferred
therapeutic regimen for the patient.
        To detect the presence of persephin or neurturin in
a patient, a sample is obtained from the patient. The
25 sample can be a tissue biopsy sample or a sample of
blood, plasma, serum, CSF or the like. Neurturin is
expressed in a wide variety of tissues as shown in
example 9 and it is believed that persephin as well is
secreted in a number of tissues. Thus, samples for
30 detecting persephin or neurturin can be taken from any
tissues expressing the particular growth factor. When
assessing peripheral levels of persephin or neurturin, it
is preferred that the sample be a sample of blood, plasma
or serum. When assessing the levels of persephin or
35 neurturin in the central nervous system a preferred
sample is a sample obtained from cerebrospinal fluid.

62

In some instances it is desirable to determine
whether the persephin or neurturin gene is intact in the
patient or in a tissue or cell line within the patient.
By an intact persephin or neurturin gene it is meant that
5 there are no alterations in the gene such as point
mutations, deletions, insertions, chromosomal breakage,
chromosomal rearrangements and the like wherein such
alteration might alter production of persephin or
neurturin or alter its biological activity, stability or
10 the like to lead to disease processes or susceptibility
to cellular degenerative conditions. Conversely, by a
non-intact persephin or neurturin gene it is meant that
such alterations are present. Thus, in one embodiment of
the present invention a method is provided for detecting
15 and characterizing any alterations in the persephin or
neurturin gene. The method comprises providing an
oligonucleotide that contains the persephin or neurturin
cDNA, genomic DNA or a fragment thereof or a derivative
thereof. By a derivative of an oligonucleotide, it is
20 meant that the derived oligonu: eotide is substantially
the same as the sequence from which it is derived in that
the derived sequence has sufficient sequence
complementarily to the sequence from which it is derived
to hybridize to the persephin or neurturin gene. The
25 derived nucleotide sequence is not necessarily physically
derived from the nucleotide sequence, but may be
generated in any manner including for example, chemical
synthesis or DNA replication or reverse transcription or
transcription.
30   Typically, patient genomic DNA is isolated from a
cell sample from the patient and digested with one or
more restriction endonucleases such as, for example, TaqI
and AluI. Using the Southern blot protocol, which is
well known in the art, this assay determines whether a
35 patient or a particular tissue in a patient has an intact

63

persephin or neurturin gene or a persephin or neurturin
gene abnormality.

Hybridization to the persephin or neurturin gene
would involve denaturing the chromosomal DNA to obtain a
5   single-stranded DNA; contacting the single-stranded DNA
with a gene probe associated with the persephin or
neurturin gene sequence; and identifying the hybridized
DNA-probe to detect chromosomal DNA containing at least a
portion of the human persephin or neurturin gene.
10   The term "probe" as used herein refers to a
structure comprised of a polynucleotide which forms a
hybrid structure with a target sequence, due to
complementarity of probe sequence with a sequence in the
target region.   Oligomers suitable for use as probes may
15   contain a minimum of about 8-12 contiguous nucleotides
which are complementary to the targeted sequence and
preferably a minimum of about 20.

The persephin or neurturin gene probes of the
present invention can be DNA or RNA oligonucleotides and
20   can be made by any method known in the art such as, for
example, excision, transcription or chemical synthesis.
Probes may be labelled with any detectable label known in
the art such as, for example, radioactive or fluorescent
labels or enzymatic marker.   Labeling of the probe can be
25   accomplished by any method known in the art such as by
PCR, random priming, end labelling, nick translation or
the like.   One skilled in the art will also recognize
that other methods not employing a labelled probe can be
used to determine the hybridization.   Examples of methods
30   that can be used for detecting hybridization include
Southern blotting, fluorescence in situ hybridization,
and single-strand conformation polymorphism with PCR
amplification.

Hybridization is typically carried out at 25-45°C,
35   more preferably at 32-40°C and more preferably at 37-
38°C.   The time required for hybridization is from about

64

0.25 to about 96 hours, more preferably from about one to
about 72 hours, and most preferably from about 4 to about
24 hours.

Persephin or neurturin gene abnormalities can also
5 be detected by using the PCR method and primers that
flank or lie within the persephin or neurturin gene. The
PCR method is well known in the art. Briefly, this
method is performed using two oligonucleotide primers
which are capable of hybridizing to the nucleic acid
10 sequences flanking a target sequence that lies within a
persephin or neurturin gene and amplifying the target
sequence. The terms "oligonucleotide primer" as used
herein refers to a short strand of DNA or RNA ranging in
length from about 8 to about 30 bases. The upstream and
15 downstream primers are typically from about 20 to about
30 base pairs in length and hybridize to the flanking
regions for replication of the nucleotide sequence. The
polymerization is catalyzed by a DNA-polymerase in the
presence of deoxynucleotide triphosphates or nucleotide
20 analogs to produce double-stranded DNA molecules. The
double strands are then separated by any denaturing
method including physical, chemical or enzymatic.
Commonly, the method of physical denaturation is used
involving heating the nucleic acid, typically to
25 temperatures from about 80°C to 105°C for times ranging
from about 1 to about 10 minutes. The process is
repeated for the desired number of cycles.

The primers are selected to be substantially
complementary to the strand of DNA being amplified.
30 Therefore, the primers need not reflect the exact
sequence of the template, but must be sufficiently
complementary to selectively hybridize with the strand
being amplified.

After PCR amplification, the DNA sequence comprising
35 persephin or neurturin or pre-pro persephin or neurturin
or a fragment thereof is then directly sequenced and

65

analyzed by comparison of the sequence with the sequences
disclosed herein to identify alterations which might
change activity or expression levels or the like.

In another embodiment a method for detecting
5   persephin or neurturin is provided based upon an analysis
of tissue expressing the persephin gene or the neurturin
gene. Certain tissues such as those identified below in
example 9 have been found to express the neurturin gene.
It is also believed that a number of tissues will express
10 the persephin gene based upon the observations for
neurturin and the identification herein of brain and
heart as tissues expressing persephin. The method
comprises hybridizing a polynucleotide to mRNA from a
sample of tissues that normally express the persephin
15 gene or the neurturin gene. The sample is obtained from
a patient suspected of having an abnormality in the
persephin gene or the neurturin gene or in the persephin
gene or the neurturin gene of particular cells. In the
case of neurturin, the polynucleotide comprises SEQ ID
20 NO:11 or a derivative thereof or a fragment thereof. In
the case of persephin, the polynucleotide comprises SEQ
ID NO:105 or SEQ ID NO:107 or the human ortholog of
persephin or derivatives thereof or fragments thereof.

To detect the presence of mRNA encoding persephin
25 protein or neurturin protein, a sample is obtained from a
patient. The sample can be from blood or from a tissue
biopsy sample. The sample may be treated to extract the
nucleic acids contained therein. The resulting nucleic
acid from the sample is subjected to gel electrophoresis
30 or other size separation techniques.

The mRNA of the sample is contacted with a DNA
sequence serving as a probe to form hybrid duplexes. The
use of a labeled probes as discussed above allows
detection of the resulting duplex.

35   When using the cDNA encoding persephin protein or
neurturin protein or a derivative of the cDNA as a probe,

66

high stringency conditions can be used in order to prevent false positives, that is the hybridization and apparent detection of persephin or neurturin nucleotide sequences when in fact an intact and functioning

5   persephin gene or neurturin gene is not present.   When using sequences derived from the persephin or neurturin cDNA, less stringent conditions could be used, however, this would be a less preferred approach because of the likelihood of false positives.   The stringency of

10 hybridization is determined by a number of factors during hybridization and during the washing procedure, including temperature, ionic strength, length of time and concentration of formamide.   These factors are outlined in, for example, Sambrook et al. (Sambrook, et al., 1989,

15 *supra*).

In order to increase the sensitivity of the detection in a sample of mRNA encoding the persephin protein or neurturin protein, the technique of reverse transcription/polymerization chain reaction (RT/PCR) can

20 be used to amplify cDNA transcribed from mRNA encoding the persephin protein or the neurturin protein.   The method of RT/PCR is well known in the art (see example 9 and figure 6 below).

The RT/PCR method can be performed as follows.

25 Total cellular RNA is isolated by, for example, the standard guanidium isothiocyanate method and the total RNA is reverse transcribed.   The reverse transcription method involves synthesis of DNA on a template of RNA using a reverse transcriptase enzyme and a 3' end primer.

30 Typically, the primer contains an oligo(dT) sequence. The cDNA thus produced is then amplified using the PCR method and persephin specific primers or neurturin specific primers.   (Belyavsky et al, *Nucl Acid Res* 17:2919-2932, 1989; Krug and Berger, *Methods in*

35 *Enzymology*, Academic Press, N.Y., Vol.152, pp. 316-325, 1987 which are incorporated by reference).

67

The polymerase chain reaction method is performed as described above using two oligonucleotide primers that are substantially complementary to the two flanking regions of the DNA segment to be amplified.

5   Following amplification, the PCR product is then electrophoresed and detected by ethidium bromide staining or by phosphoimaging.

The present invention further provides for methods to detect the presence of the persephin protein or the

10   neurturin protein in a sample obtained from a patient. Any method known in the art for detecting proteins can be used. Such methods include, but are not limited to immunodiffusion, immunoelectrophoresis, immunochemical methods, binder-ligand assays, immunohistochemical

15   techniques, agglutination and complement assays. (for example see *Basic and Clinical Immunology*, Sites and Terr, eds., Appleton & Lange, Norwalk, Conn. pp 217-262, 1991 which is incorporated by reference). Preferred are binder-ligand immunoassay methods including reacting

20   antibodies with an epitope or epitopes of the persephin protein or reacting antibodies with an epitope or epitopes of the neurturin protein and competitively displacing a labeled persephin protein or a labeled neurturin protein or derivative thereof.

25   As used herein, a derivative of the persephin protein or a derivative of the neurturin protein is intended to include a polypeptide in which certain amino acids have been deleted or replaced or changed to modified or unusual amino acids wherein the persephin

30   derivative or the neurturin derivative is biologically equivalent to persephin or neurturin, respectively, and wherein the polypeptide derivative cross-reacts with antibodies raised against the persephin protein or the neurturin protein, respectively. By cross-reaction it is

35   meant that an antibody reacts with an antigen other than the one that induced its formation.

68

Numerous competitive and non-competitive protein
binding immunoassays are well known in the art.
Antibodies employed in such assays may be unlabeled, for
example as used in agglutination tests, or labeled for
5   use in a wide variety of assay methods. Labels that can
be used include radionuclides, enzymes, fluorescers,
chemiluminescers, enzyme substrates or co-factors, enzyme
inhibitors, particles, dyes and the like for use in
radioimmunoassay (RIA), enzyme immunoassays, e.g.,
10 enzyme-linked immunosorbent assay (ELISA), fluorescent
immunoassays and the like.

Polyclonal or monoclonal antibodies to the persephin
protein and to the neurturin protein or an epitope
thereof can be made for use in immunoassays by any of a
15 number of methods known in the art. By epitope reference
is made to an antigenic determinant of a polypeptide. An
epitope could comprise 3 amino acids in a spacial
conformation which is unique to the epitope. Generally
an epitope consists of at least 5 such amino acids.
20 Methods of determining the spatial conformation of amino
acids are known in the art, and include, for example, x-
ray crystallography and 2 dimensional nuclear magnetic
resonance.

One approach for preparing antibodies to a protein
25 is the selection and preparation of an amino acid
sequence of all or part of the protein, chemically
synthesizing the sequence and injecting it into an
appropriate animal, usually a rabbit or a mouse (See
Example 10).
30   Oligopeptides can be selected as candidates for the
production of an antibody to the persephin protein or for
the production of an antibody to the neurturin protein
based upon the oligopeptides lying in hydrophilic
regions, which are thus likely to be exposed in the
35 mature protein.

69

Antibodies to persephin or to neurturin can also be
raised against oligopeptides that include one or more of
the conserved regions identified herein such that the
antibody can cross-react with other family members. Such
5   antibodies can be used to identify and isolate the other
family members.

Methods for preparation of the persephin protein or
to the neurturin protein or an epitope thereof include,
but are not limited to chemical synthesis, recombinant
10 DNA techniques or isolation from biological samples.
Chemical synthesis of a peptide can be performed, for
example, by the classical Merrifeld method of solid phase
peptide synthesis (Merrifeld, *J Am Chem Soc 85*:2149, 1963
which is incorporated by reference) or the FMOC strategy
15 on a Rapid Automated Multiple Peptide Synthesis system
(DuPont Company, Wilmington, DE) (Caprino and Han, *J Org
Chem 37*:3404, 1972 which is incorporated by reference).

Polyclonal antibodies can be prepared by immunizing
rabbits or other animals by injecting antigen followed by
20 subsequent boosts at appropriate intervals. The animals
are bled and sera assayed against purified persephin
protein or purified neurturin protein usually by ELISA or
by bioassay based upon the ability to block the action of
persephin or neurturin, as the case may be. When using
25 avian species, e.g. chicken, turkey and the like, the
antibody can be isolated from the yolk of the egg.
Monoclonal antibodies can be prepared after the method of
Milstein and Kohler by fusing splenocytes from immunized
mice with continuously replicating tumor cells such as
30 myeloma or lymphoma cells. (Milstein and Kohler *Nature
256*:495-497, 1975; Gulfre and Milstein, *Methods in
Enzymology: Immunochemical Techniques 73*:1-46, Langone
and Banatis eds., Academic Press, 1981 which are
incorporated by reference). The hybridoma cells so
35 formed are then cloned by limiting dilution methods and

70

supernates assayed for antibody production by ELISA, RIA
or bioassay.

    The unique ability of antibodies to recognize and
specifically bind to target proteins provides an approach
5 for treating an over expression of the protein. Thus,
another aspect of the present invention provides for a
method for preventing or treating diseases involving over
expression of the persephin protein or the neurturin
protein by treatment of a patient with specific
10 antibodies to the persephin protein or to the neurturin
protein, respectively.

    Specific antibodies, either polyclonal or
monoclonal, to the persephin protein or to the neurturin
protein can be produced by any suitable method known in
15 the art as discussed above. For example, murine or human
monoclonal antibodies can be produced by hybridoma
technology or, alternatively, the persephin protein or
the neurturin protein, or an immunologically active
fragment thereof, or an anti-idiotypic antibody, or
20 fragment thereof can be administered to an animal to
elicit the production of antibodies capable of
recognizing and binding to the persephin protein or to
the neurturin protein. Such antibodies can be from any
class of antibodies including, but not limited to IgG,
25 IgA, IgM, IgD, and IgE or in the case of avian species,
IgY and from any subclass of antibodies.

    Preferred embodiments of the invention are described
in the following examples. Other embodiments within the
scope of the claims herein will be apparent to one
30 skilled in the art from consideration of the specifica-
tion or practice of the invention as disclosed herein.
It is intended that the specification, together with the
examples, be considered exemplary only, with the scope
and spirit of the invention being indicated by the claims
35 which follow the examples.

71

Example 1

This example illustrates the isolation and purification of neurturin from CHO cell conditioned medium.

5 <u>Preparation of CHO cell conditioned medium:</u>

A derivative of DG44 Chinese hamster ovary cells, DG44CHO-pHSP-NGFI-B (CHO) cells, was used (Day et al, *J Biol Chem* 265:15253-15260, 1990 which is incorporated by reference). As noted above, the inventors have also

10 obtained neurturin in partially purified form from other derivatives of DG44 Chinese hamster ovary cells. The CHO cells were maintained in 20 ml medium containing minimum essential medium (MEM) alpha (Gibco-BRL No. 12561, Gaithersburg, MD) containing 10% fetal calf serum

15 (Hyclone Laboratories, Logan, UT), 2 mM l-glutamine, 100 U/ml penicillin, 100 μg/ml streptomycin and 25nM methotrexate using 150 cm² flasks (Corning Inc., Corning NY). For passage and expansion, medium from a confluent flask was aspirated; the cells were washed with 10 ml

20 phosphate buffered saline (PBS) containing in g/l, 0.144 KH_2PO_4, 0.795 Na_2HPO_4 and 9.00 NaCl; and the flask was then incubated for 2-3 minutes with 2 ml 0.25% trypsin in PBS. Cells were then knocked off the flask surface, 8 ml of medium were added and cells were triturated several times

25 with a pipette. The cells were split 1:5 or 1:10, incubated at 37°C under an atmosphere of 5% CO_2 in air and grown to confluence for 3-4 days.

The cell culture was then expanded into 850 cm² roller bottles (Becton Dickinson, Bedford, MA). A

30 confluent 150 cm² flask was trypsinized and seeded into one roller bottle containing 240 ml of the above modified MEM medium without methotrexate. The pH was maintained either by blanketing the medium with 5% CO_2 in air or by preparing the medium with 25 mM HEPES pH 7.4 (Sigma, St.

35 Louis, MO). The roller bottles were rotated at 0.8-1.0

72

revolutions per minute. Cells reached confluence in 4
days.

For collecting conditioned medium, serum-free CHO
cell (SF-CHO) medium was used. SF-CHO was prepared using
5   1:1 DME/F12 base medium, which was prepared by mixing 1:1
(v/v) DMEM (Gibco-BRL product No. 11965, Gibco-BRL,
Gaithersburg, MD) with Ham's F12 (Gibco-BRL product No.
11765). The final SF-CHO medium contained 15 mM HEPES pH
7.4 (Sigma, St. Louis, MO), 0.5 mg/ml bovine serum
10   albumin (BSA, Sigma, St. Louis MO), 25 µg/ml heparin,
(Sigma, St. Louis, MO), 1X insulin-transferrin-selenite
supplement (bovine insulin, 5 µg/ml; human transferrin, 5
µg/ml; sodium selenite, 5 ng/ml; Sigma, St. Louis, MO), 2
mM 1-glutamine, 100 U/ml penicillin, and 100 µg/ml
15   streptomycin. The medium from the confluent roller
bottles was removed and the cells washed once with 30 ml
SF-CHO medium to remove serum proteins. Cells were then
incubated at 37°C for 16-24 hrs in 80 ml SF-CHO medium to
further remove serum proteins. The 80 ml medium was
20   removed and discarded. A volume of 120 ml of SF-CHO
medium was added to the flask and the cells incubated at
37°C. Every 48 hrs thereafter, 120 ml was collected and
replaced with the same volume of SF-CHO medium.

Collected media was pooled and centrifuged at 4°C in
25   polypropylene conical tubes to remove cellular debris and
the supernatant stored at -70°C. Media was collected 5
times over 10 days to yield a total of approximately 600
ml conditioned medium per roller bottle.

Fractions collected from the columns at each stage
30   of purification were assayed for biological activity
using the neuronal survival assay and for protein content
by the dye binding assay of Bradford (Anal Biochem 72:248
et seq., 1976 which is incorporated by reference). The
total mg of protein in the starting volume, typically 50
35   liters, of conditioned medium was determined.

73

<u>Superior Cervical Ganglion Survival Assay:</u>

The neurotrophic activity of CHO conditioned medium
starting material and at various stages of purification
was assessed using the superior cervical ganglion

5   survival assay system previously reported (Martin, et al
J of Cell Biology 106:829-844; Deckwerth and Johnson, _J
Cell Bio 123_:1207-1222, 1993 which are incorporated by
reference). Primary cultures of sympathetic neurons from
superior cervical ganglion (SCG) were prepared by

10   dissecting tissue from Day 20-21 rat embryo (E20-E21).
The SCG's were placed in Leibovitz's L15 with 1-glutamine
medium (Cat #11415-023 Gibco-BRL, Gaithersburg, MD),
digested for 30 minutes with 1 mg/ml collagenase (Cat
#4188 Worthington Biochemical, Freehold, NJ) in

15   Leibovitz's L15 medium at 37°C, followed by a 30 minute
digestion in trypsin-lyophilized & irradiated (Type
TRLVMF Cat #4454 Worthington Biochemical, Freehold, NJ)
which was resuspended in modified Hanks' Balanced Salt
Solution (Cat #H-8389 Sigma Chemical Co., St. Louis, MO).

20   The digestion was stopped using AM50 which contains
Minimum Essential Medium with Earle's salts and without
1-glutamine (Cat #11090-016 Gibco-BRL), 10% fetal calf
serum (Cat #1115 Hyclone Laboratories, Logan, UT), 2mM 1-
glutamine (Cat #G5763 Sigma Chemical Co., St. Louis, MO),

25   20 µM FuDr (F-0503 Sigma Chemical Co., St. Louis, MO), 20
µM Uridine (Cat #3003 Sigma Chemical Co., St. Louis, MO),
100 U/ml penicillin, 100 µg/ml Streptomycin, and 50 ng/ml
2.5 S NGF. The cells were dissociated into a suspension
of single cells using a silanized and flame-polished

30   Pasteur pipet. After filtration of the suspension
through a nitex filter (size 3-20/14, Tetko Inc.,
Elmsford, NY), the cells were placed in AM50 medium as
above and preplated on a 100 mm Falcon or Primaria
culture dish (Becton Dickinson Labware, Lincoln Park, NJ)

35   to reduce the number of non-neuronal cells. After 2
hours, the medium containing the unattached neuronal

74

cells was removed from these dishes and triturated again
through a silanized and flame-polished Pasteur pipet.
The single cell suspension was plated on 24-well tissue
culture plates (Costar, Wilmington, MA) that have been
5   previously coated with a double layer of collagen, one
layer of collagen that had been ammoniated and a second
layer of collagen that had been air dried. They were
allowed to attach for 30 minutes to 2 hours. A specific
number of viable cells, usually about 1200 to about 3000
10 total cells per well, or a specific percentage of the
ganglion, usually 25% of the cells obtained per ganglion
were plated into each well. When cell counts were to be
performed they were placed in the 24-well dishes as
stated above or alternatively, on 2-well chamber slides
15 (Nunc, Naperville, IL). Cultures were then incubated for
5-6 days at 37° in AM50 medium in a 5% CO_2/95% air
atmosphere. The death of the cultured neurons was
induced by exchanging the medium with medium without NGF
and with 0.05% goat anti-NGF (final titer in the wells is
20 1:10). This NGF-deprivation results in death of the
neurons over a period of 24-72 hours. Aliquots of
partially purified or purified factor, or appropriate
controls, were added to the cultures at the time of NGF
removal to determine the ability to prevent the neuronal
25 death.
        Evaluation of the ability of column fraction, gel
eluates, or purified factor to prevent neuronal death was
by visual inspection of cultures under phase contrast
microscopy. Viable neurons remained phase bright with
30 intact neurities, whereas dead neurons were shrunken,
phase dark, had irregular membranes and neurites were
fragmented (Figure 3). Where precise quantitation of
neuronal survival was required, the cultures were fixed
in 4% paraformaldehyde or 10% Formalin in PBS, and
35 stained with crystal violet solution, (Huntoon Formula
Harleco E.M. Diagnostics Systems, Gibbstown, NJ). When

75

using 24 well dishes, 1 μl crystal violet solution was
added to each well containing 10% formalin and the cells
were counted using a phase contrast microscope. If the
2-well chamber slides were used, the cultures were fixed,
5   stained with crystal violet, destained with water,
dehydrated in increasing ethanol concentrations to
toluene, and mounted in a toluene-based mounting
solution. Neurons were scored as viable if they had a
clear nucleolus and nuclei and were clearly stained with
10   crystal violet.
        The neuronal death at 72 hours in shown in Figure
3B. Also shown are (A) the positive control cells
maintained with nerve growth factor and (C) the cells
treated with anti-NGF and neurturin (approximately 3
15   ng/ml) showing survival of neurons.
        Activity was quantitated by calculation of a
"survival unit". The total survival units in a sample
were defined as the minimal volume of an aliquot of the
sample which produced maximal survival divided into the
20   total volume of that sample. Specific activity was
calculated as the survival units divided by the mg total
protein.
        Survival units were determined in an assay using
approximately 1200 viable neurons in a 0.5 ml culture
25   assay and a culture period of 48 hours following addition
of the fraction. Survival was assessed visually after
the 48 hours. Intrinsic activity as shown in Figure 4
was determined in an assay using approximately 2700
neurons and a culture period of 72 hours. Survival was
30   assessed by fixing the neurons and counting the number of
surviving neurons. Because the stability, as assessed by
half-life of activity, for neurturin decreases as the
number of neurons increases, the intrinsic activity
measurement would be expected to be lower than that
35   predicted by Specific Activity determinations. The
intrinsic activity measurement would also be expected to

76

be lower than that predicted by specific activity because
the survival was measured after 72 hours instead of 48
hours.

To ensure the reproducibility of these activity unit

5   assays, it was necessary to plate the primary neuronal
cultures at reproducible cell densities, as the stability
of the activity decreases significantly with increasing
neuronal density. The range of cell densities was from
about 1200 to about 2700 cells per well. The presence of

10   soluble heparin in the assay medium had no effect on the
short-term (~3 days) stability of the survival activity.
Purification of Neurturin:

Pooled conditioned medium was filtered through 0.2
μl pore bottle-top filters (cellulose acetate membrane,

15   Corning Inc., Corning, NY). Typically 50 liters of
conditioned medium was used and processed in 25 liter
batches. Each 25 liter batch was introduced at a rate of
20 ml/min onto a 5 x 5 cm column containing 100 ml
heparin-agarose (Sigma, St. Louis, MO) equilibrated with

20   25 mM HEPES, pH 7.4 buffer with 150 mM NaCl. The column
was then washed with approximately 1000 ml 25 mM HEPES,
pH 7.4 buffer containing 0.5 M NaCl at 20 ml/min and the
activity was then eluted with 25 mM HEPES, pH 7.4 buffer
containing 1.0 M NaCl. After switching to the 1.0M NaCL

25   elution buffer, the first 50 ml of buffer was discarded
and, thereafter, one 300 ml fraction was collected.

Pooled material eluted from the Heparin-agarose
column was then diluted 1:1 (v/v) with 25 mM HEPES, pH
7.4 buffer containing 0.04% TWEEN 20 to a NaCl

30   concentration of 0.5 M and introduced into a 1.5 cm x 9
cm column containing 16 ml SP SEPHAROSE® High Performance
ion exchange resin (Pharmacia, Piscataway, NJ)
equilibrated in 25 mM HEPES 7.4 containing 0.5 M NaCl and
0.02% TWEEN 20. The column was then washed with 160 ml

35   25 mM HEPES, pH 7.4 buffer containing 0.5 M NaCl and
0.02% TWEEN 20 and the activity was eluted with 25 mM

77

HEPES, pH 7.4 buffer containing 1.0 M NaCl and 0.02%
TWEEN 20 at a flow rate of 2 ml/min. One 50 ml fraction
was collected after the first 7 ml of eluate from the
column.

5   Material eluted from the SP SEPHAROSE® column was
fractionated using fast protein liquid chromatography
(FPLC) on a Chelating Superose HR 10/2 column charged
with Cu⁺⁺ (Pharmacia, Piscataway, NJ). The column had
been prepared by washing with 10 ml water, charging with
10   3 ml of 2.5 mg/ml CuSO₄·5H₂O, washing with 10 ml water,
and equilibrating with 10 ml of 25 mM HEPES pH 7.4 buffer
containing 1.0 M NaCl and 0.02% TWEEN 20. The eluate was
introduced into the column in 25 mM HEPES, pH 7.4 buffer
containing 1.0 M NaCl at a rate of 1.0 ml/min. The bound
15   proteins were eluted with a linear gradient of increasing
glycine concentration (0-300 mM) in 25 mM HEPES, pH 7.4
buffer containing 1.0 M NaCl at a rate of 1.0 ml/min.
The gradient was produced by a Pharmacia FPLC system
using an LCC-500 controller and P-500 pumps to establish
20   a 0-300 mM glycine gradient in 40 ml at 1.0 ml/min, thus
increasing the gradient by 7.5 mM glucine per min. One
ml fractions were collected and assayed for SCG survival
promotion. Peak activity was observed in fractions 17-
20, i.e. 17-20 min or ml from the start of the gradient.

25   Absorbance measurements at 280 nM by an in-line UV
monitor indicated that most proteins eluted prior to the
survival activity in fractions 17-20. Thus, significant
purification was achieved at this step. A 25 kD band co-
purified with the survival activity.

30   The combined eluted fractions from the Cu⁺⁺ superose
column were diluted to 0.45 M NaCl using 25 mM HEPES pH
7.4 buffer containing 0.02% TWEEN 20 and introduced into
a Mono S HR 5/5 cation exchange column (Pharmacia,
Piscataway, NJ) for further FPLC purification. The
35   column had been equilibrated with 25 mM HEPES pH 7.4
buffer containing 0.45 M NaCl containing 0.02% TWEEN 20.

78

Bound proteins were eluted with a linear gradient of increasing NaCl concentration (0.45-1.0 M). The gradient was produced as described above from 0.45 M - 1.0 M NaCl in 35 mls at 1.0 ml/min, thus increasing concentration at

5   0.0157 M per ml or min. Thirteen 1.0 ml fractions (fractions 1-13) were collected followed by 44 0.5 ml fractions (fractions 14-53). Peak activity in SCG assay was in fractions 26-29. Each fraction was assayed in the SCG survival assay over a range of volumes of from 0.1 to

10 1.0 µl per 0.5 ml culture medium.

One percent (5 µl) of each fraction was loaded onto a non-reducing, 14% SDS polyacrylamide gel and electrophoresed for 750 V-hr at 25°C. Proteins were visualized by silver stain. The results are shown in

15 Figure 2. Markers shown in lane M on the gel represent 20 ng of Bovine serum albumin, carbonic anhydrase, B-lactoglobulin, and lysozyme in the order of descending molecular weight.

A 25 kD band appeared in fractions 25-30, a 28 kD

20 protein elutes earlier in the gradient and an 18 kD elutes later in the gradient. Figure 2 illustrates the survival activity in each of the fractions. The survival activity is noted to correspond with the presence and apparent intensity of the 25 kD protein in fractions 25-

25 30.

To demonstrate that the 25 kD band was responsible for survival promoting activity, the 25 kD protein was eluted from the polyacrylamide gel after electrophoresis and assayed for survival activity in the SCG assay.

30 After electrophoresis of 150 µl of the SP SEPHAROSE® 1.0 M NaCl fraction in one lane of a non-reducing 14% SDS-polyacrylamide gel as above, the lane was cut into 12 slices and each slice was crushed and eluted by diffusion with rocking in buffer containing 25 mM HEPES, pH 7.4,

35 0.5 M NaCl, 0.02% Tween-20 for 18 hr at 25°C. BSA was added to the eluate to a final concentration of 200 µg/ml

and the eluate was filtered through a 0.45 micron filter
to remove acrylamide gel fragments. The filtrate was
then added to a SP SEPHAROSE® column to concentrate and
purify the sample. Before eluting the sample, the column
5 was washed once in 400 µl 25 mM HEPES, pH 7.4 buffer
containing 0.5 M NaCl, 0.02% Tween-20 and 200 µg BSA per
ml and once in 400 µl 25 mM HEPES, pH 7.4 buffer
containing 0.02% Tween-20 and 200 µg BSA per ml. The
column was then washed again in 400 µl of 25 mM HEPES, pH
10 7.4 buffer containing 0.5 M NaCl, 0.02% TWEEN 20 and 200
ug BSA per ml. The sample was eluted with 25 mM HEPES,
pH 7.4 buffer containing 1.0 M NaCl, 0.02% Tween-20 and
200 µg BSA per ml. Samples were then analyzed for
survival activity. Only the slice corresponding to the
15 25 kD band showed evidence of survival activity. The 25
kD protein purified from CHO cell conditioned media is
believed to be a homodimer.

The yield from the purification above was typically
1-1.5 µg from 50 liters of CHO cell conditioned medium.
20 Overall recovery is estimated to be 10-30%, resulting in
a purification of approximately 390,000 fold.

Example 2

This example illustrates the characterization of
25 neurturin and several members of the TGF-β family of
growth factors in the SCG assay and the lack of cross
reactivity of anti-GDNF antibodies with neurturin.

The SCG assay of the purified protein indicated that
the factor is maximally active at a concentration of
30 approximately 3 ng/ml or approximately 100 pM and the EC_{50}
was approximately 1.5 ng/ml or approximately 50 pM in the
expected range for a diffusible peptide growth factor
(Figure 4).

Several members of the TGF-β family influence
35 neuropeptide gene expression in sympathetic neurons,
while others promote survival of different neuronal

80

populations. Neurturin, which is a distant member of
this family of proteins, is capable of promoting
virtually complete survival of sympathetic neurons for 3
days. In addition, further culturing of the SCG cells
5   revealed that neurturin could continue to maintain these
neurons for at least 10 days after withdrawal of NGF.

      We tested several other members of the TGF-β family
for their ability to promote survival in the SCG assay
including TGF-β1, activin, BMP-2, BMP-4, BMP-6 and GDNF.
10 Of these factors, only GDNF had survival promoting
activity, however, the activity of GDNF was much less
potent than neurturin in this activity showing an EC_{50} of
2-4 nM in the 3-day survival assay. The GDNF tested in
this assay was rhGDNF produced in *E. Coli* obtained from
15 Prepro Tech, Inc., Rocky Hill, N.J. The duration of
action of GDNF was also less than that of neurturin
inasmuch as the ability of GDNF (50 ng/ml) to maintain
survival longer than 3 day was substantially diminished.
These experiments suggest the possibility that GDNF is a
20 weak agonist for the neurturin receptor. Furthermore,
the inability of activin and BMP-2 to promote survival,
in contrast to their strong induction of transmitter-
related gene expression in these neurons (Fann and
Paterson, *Int J Dev Neurosci* 13:317-330, 1995; Fann and
25 Patterson, *J Neurochem* 61:1349-1355, 1993) suggests that
they signal through alternate receptors or signal
transduction pathways.

      To determine the cross-reactivity of anti-GDNF
antibodies with partially purified neurturin, SCG
30 neurons, that had been dissected and plated as described
in Example 1 were treated on Day 6 with 1 ng/ml, 3 ng/ml,
10 ng/ml, or 30 ng/ml GDNF (Prepro Tech, Inc, Rocky Hill,
N.J.) in the presence of anti-NGF alone, or in the
presence of anti-NGF and anti-GDNF (goat IgG antibody to
35 *E. coli*-derived rhGDNF, R & D Systems, Minneapolis,
Minn). A partially purified 1.0 M SP Sepharose fraction

81

of neurturin was used in the assay at the approximate
concentrations of 375 pg/ml, 750 pg/ml, 1.5 ng/ml and 3
ng/ml. This fraction was tested in the presence of
anti-NGF alone, and in the presence of anti-NGF and
5   anti-GDNF. The anti-GDNF antibody blocked the survival
promoting activity of GDNF at a concentration up to 30
ng/ml, but did not block the survival promoting activity
of neurturin.

10   Example 3
        This example illustrates the effect of neurturin on
sensory neurons in a nodose ganglion survival assay.
        CHO cell conditioned media that had been partially
purified on the SP Sepharose column was assayed for
15 neurotrophic activity on sensory neurons using nodose
ganglia. The survival assay is a modification of that
previously reported above for superior cervical ganglia.
Primary dissociated cultures of nodose ganglia were
prepared by dissecting tissue from E18 Sprague Dawley rat
20 pups. The nodose ganglia were placed in Leibovitz's L15
with 2 mM 1-glutamine (Cat# 11415-023, GIBCO-BRL.
Gaithersburg, MD) as the tissues was dissected, digested
for 30 min with 1 mg/ml collagenase (Cat#4188,
Worthington Biochemical, Freehold, New Jersey) in
25 Leibovitz's L15 medium at 37°C, followed by 30 min
digestion in trypsin (lyophilized and irradiated, type
TRLVMF, Cat #4454 Worthington Biochemical, Freehold, NJ),
and resuspension to a final concentration of 0.25% in
modified Hank's Balanced Salt Solution (Cat#H8389, Sigma
30 Chemical Co., St. Louis, Mo). The digestion was stopped
using AMO-BDNF100, a medium containing Minimum Essential
Medium with Earle's salts and without 1-glutamine
(#11090-016 GIBCO-BRL), 10% fetal Calf Serum (Cat#1115,
Hyclone Laboratories, Logan, UT), 2 mM 1-glutamine
35 (Cat#G5763 Sigma Chemical Co., St. Louis, Mo.), 20 μM
FuDr (F-0503, Sigma Chemical Co.), 20 μM Uridine (Cat

82

#3003, Sigma Chemical Co., St. Louis, Mo.) 100 U/ml
penicillin, 100 µg/ml Streptomycin, and 100 ng Brain
Derived Neurotropic Factor (BDNF, Amgen, Thousand Oaks,
CA). The cells were dissociated into a suspension of
5   single cells using a silanized and flame-polished Pasteur
pipet in the AMO-BDNF100 medium, and preplated on a 100
mm Falcon or Primaria culture dish (Becton Dickinson
Labware, Lincoln Park, NJ) to remove non-neuronal cells.
After 2 hours, the medium containing the unattached
10 neuronal cells was removed from these dishes and
triturated again through a silanized and flame-polished
Pasteur pipet. The single cell suspension was plated on
24-well tissue culture plates (Costar, Wilmington, MA)
that have been previously coated with a double layer of
15 collagen, one layer of which had been ammoniated and a
second layer that had been air dried. Ganglia from ten
E18 rat embryos were dissociated into 2.5 mls of media
and 100 µl of this suspension was added to each well.
The cells were allowed to attach for 30 min in a 37°C
20 incubator with 5% CO2/95% air. The wells were fed with
AMO-BDNF100 media overnight.

The next day the cells were washed 3 times for 20
min each time with AMO medium containing no BDNF. The
wells were fed with 0.5 ml of this media alone or this
25 media containing either 50 ng/ml NGF, 100 ng/ml BDNF
(Amgen, Thousand Oaks, CA), 100 ng/ml GDNF (Prepro Tech,
Inc., Rocky Hill, N.J) or 3 ng/ml Neurturin. The cells
were incubated at 37°C in a 5% CO_2/95% air incubator for 3
days, fixed with 10% formalin, stained with crystal
30 violet (1 µl/ml 10% formalin) and counted. Survival was
ascertained as noted previously.

The neuronal Death at 72 hours is shown in Figure
10. Neuronal survival of nodose neurons cultured in
BDNF has been previously reported (Thaler et al, *Develop
35 Biol 161*:338-344, 1994 which is incorporated by
reference). This was used as the standard for survival

83

for these neurons and given the value of 100% survival.
Nodose ganglia that had no trophic support (AMO) showed
20%-30% survival, as did neurons that were cultured in
the presence of 50 ng/ml NGF. Neurons cultured in the

5   presence of 3 ng/ml neurturin and absence of BDNF showed
survival similar to those neurons cultured in the
presence of BDNF (100 ng/ml). GDNF at a concentration of
100 ng/ml promoted greater survival of nodose neurons
than did BDNF (100 ng/ml). Similar findings with GDNF

10   were recently reported for sensory neurons from chicken
(Ebendal, T. et al, *J Neurosci Res 40*:276-284 1995 which
is incorporated by reference).

Example 4

15   This example illustrates the determination of
partial amino acid sequences of neurturin isolated from
CHO cell conditioned medium.
         To obtain N-terminal amino acid sequence from a
purified preparation of approximately 1 µg of neurturin,

20   the Mono S fractions 26-29 containing the peak of
activity were concentrated to 25 µl by centrifuge
ultrafiltration in a microcon-3 concentrators (Amicon,
Inc., Beverley, MA) and loaded onto a non-reducing 14%
SDS polyacrylamide gel. After electrophoretic

25   separation, proteins were electroblotted to a PVDF
membrane (Bio-Rad, Hercules, CA) and stained with 0.1%
Coomassie Blue. The 25 kD band was excised and inserted
into the reaction cartridge of an automated sequencer
(Model 476, Applied Biosystems (Foster City, CA).

30   Phenylthiohydantoin-amino acid (PTH-aa) recovery in the
first 2-3 cycles of automated sequencing by Edman
degradation indicated a sequencing yield of 4 pmoles,
which was approximately 10% of the estimated amount of
protein loaded on the SDS gel.

35   Two N-terminal sequencing runs were performed from
two 50 liter purification preparations. In the first

84

run, 1 µg of protein in 3 pooled fractions of 1.5 ml
total volume were concentrated to 25 µl and
electroblotted at 100V for 2 hrs at 25°C using an
electroblot buffer of 10 mM CAPS pH 11.0 buffer (Sigma,
5   St. Louis, MO) containing 5% methanol. The amino acid
sequence was obtained from 13 cycles of Edman degradation,
and the sequencing yield was 4 pmoles as above.

        In the second run, 1.5 µg of protein in 4 pooled
fractions of 2.0 ml total volume were concentrated to 25
10 µl and electroblotted at 36V for 12 hours at 4°C using an
electroblot buffer of 25 mM Tris, 192 mM glycine, 0.04%
SDS and 17% MeOH. Sequencing yield was 15 pmoles and the
sequence after 16 cycles was SGARPXGLRELEVSVS (SEQ ID
NO:3). The sequence obtained after 16 cycles
15 corresponded to the shorter sequence obtained in the
first run. Definite assignments could not be made at 3
of the amino acid residues in the sequence (residues 1, 6
and 11 from the N-terminal). A search of protein
databases did not detect any significantly homologous
20 sequences, suggesting that the purified factor was a
novel protein.

        This initial N-terminal amino acid sequence data did
not enable the isolation of cDNA clones using degenerate
oligonucleotides as PCR primers or probes for screening
25 libraries. To facilitate these approaches, additional
protein was purified in order to obtain internal amino
acid sequence from proteolytic fragments. To obtain
internal amino acid sequence from neurturin, an
additional 50 liters of CHO cell conditioned medium was
30 purified using only the first 3 chromatographic steps as
outlined above, except that the gradient used to elute
the Cu++ Chelating Superose column was as follows: 0-60
mM glycine (4 ml), 60mM glycine (10ml), 60-300 mM glycine
(32 ml). Fractions No. 20-23 containing neurturin were
35 concentrated to 25 µl by ultrafiltration (Amicon microcon
3, Amicon, Beverley, MA) and loaded on a non-reducing SDS

85

polyacrylamide gel. After electrophoresis, the gel was
stained with Coomassie blue and the 25 kD neurturin band
was excised. Neurturin was digested in the gel slice
with endoproteinase Lys-C, and the eluted proteolytic
5   fragments were purified by reverse phase HPLC. Only one
peak was observed upon HPLC separation of the eluted
peptides, which yielded amino acid sequence information
for 23 cycles at the 1 pmole signal level using the
automated sequencer, (internal fragment P2, SEQ ID NO:5).
10   Amino acid analysis performed on 10% of the above
sample before subjecting it to digestion had indicated
that 150 pmoles of protein were present in the gel slice,
consisting of 7.6% lysine and 19.5% arginine. The single
low level peak from the Lys-C digestion suggested that
15 the digestion and elution of peptides were inefficient.
The same gel slice was redigested with trypsin and the
eluted peptides separated by HPLC. Two peaks were
observed on HPLC, resulting in the elucidation of two
additional 10 residue amino acid sequences (4-5 pmole
20 signal level, internal fragment P1, SEQ ID NO:4 and
internal fragment P3, SEQ ID NO:6) that were distinct
from the N-terminal and previous internal amino acid
sequences. The in situ digestion, elution and
purification of peptides, and peptide sequencing was
25 performed by the W.M. Keck Foundation Biotechnology
Resource Laboratory at Yale University according to
standard protocols for this service.

Example 5

30   The following example illustrates the isolation and
sequence analysis of mouse and human neurturin cDNA
clones.
Degenerate oligonucleotides corresponding to various
stretches of confident amino acid sequence data were
35 synthesized and used as primers in the polymerase chain
reaction (PCR) to amplify cDNA sequences from reverse

transcribed mRNA.   A forward primer (M1676;
5'-CCNACNGCNTAYGARGA, SEQ ID NO:50) corresponding to
peptide sequence P2 Xaa₁-Xaa₂-Val-Glu-Ala-Lys-Pro-Cys-Cys-
Gly-Pro-Thr-Ala-Tyr-Glu-Asp-Xaa₃-Val-Ser-Phe-Leu-Ser-Val

5   where Xaa₁ and Xaa₂ were unknown, Xaa₃ was Gln or Glu (SEQ
ID NO:5) in combination with a reverse primer (M1677;
5'-ARYTCYTGNARNGTRTGRTA (SEQ ID NO:52) corresponding to
peptide sequence P3
(Tyr-His-Thr-Leu-Gln-Glu-Leu-Ser-Ala-Arg) (SEQ ID NO:6)

10   were used to amplify a 69 nucleotide product from cDNA
templates derived from E21 rat and adult mouse brain.
The PCR parameters were: 94°C for 30 sec; 55°C for 30
sec; 72°C for 1 min for 35 cycles.   The product was
subcloned into the Bluescript KS plasmid and sequenced.

15   All nucleotide sequencing was performed using fluorescent
dye terminator technology per manufacturer's instructions
on an Applied Biosystems automated sequencer Model #373
(Applied Biosystems, Foster City, CA).   Plasmid DNA for
sequencing was prepared using the Wizard Miniprep kit

20   (Promega Corp., Madison, WI) according to the
manufacturer's instructions.   The sequence of the
amplified product correctly predicted amino acid sequence
data internal to the PCR primers.

        Primers corresponding to the amplified sequence were
25   used in combination with the degenerate primers in the
rapid amplification of cDNA ends (RACE) technique
(Frohman, M.A. Methods in Enzymology 218:340-356, 1993)
using the Marathon RACE kit (CLONTECH, Palo Alto, CA) per
the manufacturer's instructions, except that first strand

30   cDNA synthesis was carried out at 50°C using Superscript
II reverse transcriptase (Gibco-BRL).   Briefly, a double
stranded adaptor oligonucleotide was ligated to the ends
of double stranded cDNA synthesized from postnatal day 1
rat brain mRNA.   Using nested forward neurturin PCR

35   primers (M1676; 5'-CCNACNGCNTAYGARGA, SEQ ID NO:50 and
1678; 5'-GACGAGGGTCCTTCCTGGACGTACACA, SEQ ID NO:53) in

87

combination with primers to the ligated adaptor supplied
in the kit (AP1, AP2), the 3' end of the neurturin cDNA
was amplified by two successive PCR reactions (1st: M1676
and AP1, using 94°C for 30 sec, 55°C for 30 sec and 72°C
5 for 2 min for 35 cycles; 2nd: M1678 and AP2 using 94°C
for 30 sec and 68°C for 2 min for 35 cycles). A 5'
portion of the rat neurturin cDNA was obtained by two
successive PCR reactions using the linkered cDNA as
template. The 1st reaction utilized primers M1677 (SEQ
10 ID NO:52) and AP1; using 94°C for 30 sec; 55°C for 30
sec; and 72°C for 2 min for 35 cycles. The 2nd reaction
used M1679 5'-TAGCGGCTGTGTACGTCCAGGAAGGACACCTCGT (SEQ ID
NO:54) and AP2 at 94°C for 30 sec and 68°C for 2 min for
35 cycles. These reactions resulted in a truncated form
15 of the 5' end of the neurturin cDNA, apparently the
result of premature termination of the cDNA during
reverse transcription. The 5' and 3' RACE products were
subcloned into the plasmid Bluescript KS and sequenced.
The sequence of these 3' and 5' RACE products resulted in
20 a partial rat neurturin cDNA sequence of 220 nt. Primers
(#467921 5'-CAGCGACGACGCGTGCGCAAAGAGCG, SEQ ID NO:55; and
M1679 (SEQ ID NO:54) corresponding to the partial rat
cDNA sequence were used (PCR parameters 94°C for 30 sec
and 68°C for 1 min for 35 cycles) to amplify a 101
25 nucleotide PCR product from mouse genomic DNA which was
homologous to rat neurturin cDNA sequence.
        These primers were then used to obtain murine
neurturin genomic clones by amplifying gene fragments in
a mouse 129/Sv library in a P1 bacteriophage vector
30 (library screening service of Genome Systems, Inc., St.
Louis, MO). A 1.6 kb Nco I fragment from this P1 clone
containing the neurturin gene was identified by
hybridization with primer (#465782;
5'-TAYGARGACGAGGTGTCCTTCCTGGACGTACACAGCCGCTAYCAYAC, SEQ
35 ID NO:56). This Nco I fragment was sequenced and found
to contain a stretch of coding sequence corresponding to

88

the N-terminal and internal amino acid sequences obtained
from sequencing the active protein isolated from CHO cell
conditioned media. Beginning at the N-terminal amino
acid sequence of the purified protein, this nucleotide

5   sequence encodes a 100 amino acid protein with a
predicted molecular mass of 11.5 kD. A search of protein
and nucleic acid databases identified neurturin as a
novel protein that is approximately 40% identical to
glial derived neurotrophic factor (GDNF). GDNF was

10 purified and cloned as a factor which promotes the
survival of midbrain dopaminergic neurons and is a
distantly related member of the TGF-ß superfamily, which
now includes more than 25 different genes that possess a
wide variety of proliferative and differentiative

15 activities. Although GDNF is less than 20% identical to
any other member of the TGF-ß family, it contains the 7
cysteine residues which are conserved across the entire
family and believed to be the basis of a conserved
cysteine knot structure observed in the crystal structure

20 determination of TGF-ß2. Neurturin also contains these 7
cysteine residues, but like GDNF is less than 20%
homologous to any other member of the TGF-ß family.
Thus, neurturin and GDNF appear to represent a subfamily
of growth factors which have significantly diverged from

25 the rest of the TGF-ß superfamily.
          To determine the sequence of the full length mouse
neurturin cDNA, 5' and 3' RACE PCR was performed as
above for the rat, using nested primers predicted from
the mouse genomic sequence and cDNA from neonatal mouse

30 brain. The 1st reaction for the 3' end used primers:
M1777 5'-GCGGCCATCCGCATCTACGACCGGG (SEQ ID NO:57) and AP1
at 94°C for 30 sec; 65°C for 15 sec; and 68°C for 2 min
for 35 cycles. The 2nd reaction used primer #467921 (SEQ
ID NO:55) and AP2 at 94°C for 30 sec; 65°C for 15 sec;

35 and 68°C for 2 min for 20 cycles. The 5' end was
obtained using for the 1st reaction primer M1759,

89

5'-CRTAGGCCGTCGGGCGRCARCACGGGT (SEQ ID NO:58) and AP1 at
94°C for 30 sec; 65°C for 15 sec; and 68°C for 2 min for
35 cycles. The 2nd reaction used primer M1785,
5'-GCGCCGAAGGCCCAGGTCGTAGATGCG (SEQ ID NO:59) and AP2 at
5   94°C for 30 sec; 65°C for 15 sec; and 68°C for 2 min for
20 cycles. Both sets of PCR reactions included 5% DMSO.
The 5' and 3' mouse RACE products were subcloned into the
plasmid Bluescript KS and sequenced. Using the sequence
of RACE products, a 1.0 kb mouse neurturin cDNA sequence
10 can be assembled. This cDNA sequence contains an open
reading frame of 585 nucleotides that encodes a protein
with a molecular mass of 24 kD. This full length mouse
cDNA sequence is shown in Figure 7 (SEQ ID NO:12).
Consistent with the processing events known to occur for
15 TGF-ß family members, the 24 kD neurturin protein
contains an amino terminal 19 amino acid signal sequence
followed by a pro-domain which contains an RXXR
proteolytic processing site immediately before the
N-terminal amino acid sequence obtained when sequencing
20 the protein purified from CHO cell conditioned media.
Using these landmarks, the 11.5 kD mature neurturin
molecule is predicted to be 11.5 kD and, by analogy to
other members of the TGF-ß family, is predicted to form a
disulfide linked homodimer of 23 kD, consistent with the
25 25 kD mass of the protein purified from CHO cell
conditioned media as estimated by SDS-PAGE analysis.
     For isolation of human genomic clones, primers
(#467524; 5'-CGCTACTGCGCAGGCGCGTGCGARGCGGC, SEQ ID NO:60
and #10005, 5'-CGCCGACAGCTCTTGCAGCGTRTGGTA, SEQ ID NO:61)
30 predicted from the sequence of mouse neurturin were used
to amplify (PCR parameters: Initial denaturation at 95°C
for 1 min 30 sec followed by 94°C for 30 sec; 60°C for 15
sec; and 68°C for 60 sec for 35 cycles) a 192 nucleotide
fragment from human genomic DNA. The sequence of the PCR
35 product demonstrated that it was the human homolog of
mouse neurturin. The primers were then used to screen a

human genomic library constructed in the Pl vector
(library screening service, Genome Systems, Inc.) and two
clones containing the human neurturin genomic locus were
obtained.

5   The same strategy was used to determine the human
sequence as discussed above for the mouse sequence. An
oligo (#30152, GACCTGGGCCTGGGCTACGCGTCCGACGAG, SEQ ID
NO:62) was used as a probe in a Southern blot analysis to
identify restriction fragments of the Pl Clones which
10 contained the human neurturin coding sequence. These
restriction fragments (Eag I, Pvu II, Hind III, Kpn I)
were subcloned into the Bluescript KS plasmid and
sequenced.

         The results of subcloning and sequencing of human
15 genomic fragments were as follows. The Eag I fragment
was found to be approximately 6 kb in size with the 3'
Eag I site located 60 bp downstream from the stop codon.
The Pvu II fragment was approximately 3.5 kb in size with
the 3' Pvu II site located 250 bp downstream from the
20 stop codon. The Hind III fragment was approximately 4.8
kb in size with the 3' Hind III site located 3kb
downstream from the stop codon. The Kpn I fragment was
approximately 4.2 kb in size with the 3' Kpn I site
located 3.1 kb downstream from the stop codon.

25   The second coding exon was sequenced using these
subcloned fragments. In addition, sequence was obtained
from 250 bp flanking the 3' side of the second exon. The
sequence was also obtained from 1000 bp flanking the 5'
side of the coding exon. From these flanking sequences,
30 forward primer 30341 (5'-CTGGCGTCCCAMCAAGGGTCTTCG-3', SEQ
ID NO:71) and reverse primer 30331 (5'-
GCCAGTGGTGCCGTCGAGGCGGG-3', SEQ ID NO:72) were designed
so that the entire coding sequence of the second exon
could be amplified by PCR.

35   The first coding exon was not mapped relative to the
restriction sites above but was contained in the Eag I

91

fragment.   The sequence of this exon was obtained from
the subcloned Eag I fragment using the mouse primer
466215 (5'-GGCCCAGGATGAGGCGCTGGAAGG-3', SEQ ID NO:73),
which contains the ATG initiation codon.   Further
5   sequence of the first coding exon was obtained with
reverse primer 20215 (5'-CCACTCCACTGCCTGAWATTCWACCCC-3',
SEQ ID NO:74), designed from the sequence obtained with
primer 466215.   Forward primer 20205 (5'-
CCATGTGATTATCGACCATTCGGC-3', SEQ ID NO:75) was designed
10   from sequence obtained with primer 20215.   Primers 20205
and 20215 flank the coding sequence of the first coding
exon and can be used to amplify this coding sequence
using PCR.

                            Example 6
15   This example illustrates the preparation of
expression vectors containing neurturin cDNA.
        For expression of recombinant neurturin in mammalian
cells the neurturin vector pCMV-NTN-3-1 was constructed.
20   The 585 nucleotide open reading frame of the neurturin
cDNA was amplified by PCR using a primer containing the
first 27 nucleotides of the neurturin coding sequence
(5'-GCGACGCGTACCATGAGGCGCTGGAAGGCAGCGGCCCTG, SEQ ID
NO:63) and a primer containing the last 5 codons and the
25   stop codon (5'-GACGGATCCGCATCACACGCACGCGCACTC) (SEQ ID
NO:64) using reverse transcribed postnatal day 1 mouse
brain mRNA as template using (PCR parameters: 94°C for 30
sec; 60°C for 15 sec; and 68°C for 2 min for 35 cycles
and including 5% DMSO in the reaction).   The PCR product
30   was subcloned into the Eco RV site of BSKS and sequenced
to verify that it contained no PCR generated mutations.
The neurturin coding sequence was then excised from this
vector using Mlu I (5' end) and Bam H1 (3' end) and
inserted downstream of the CMV IE promoter/enhancer in
35   the mammalian expression vector pCB6 (Brewer, C.B.

92

Methods in C 11 Biology 43:233-245, 1994) to produce the
pCMV-NTN-3-1 vector using these sites.

For expression of recombinant protein in E. Coli,
the mature coding region of mouse neurturin was amplified
5   by PCR using a primer containing the first 7 codons of
the mature coding sequence
(5'-GACCATATGCCGGGGGCTCGGCCTTGTGG) (SEQ ID NO:65) and a
primer containing the last 5 codons and the stop codon
5'-GACGGATCCGCATCACACGCACGCGCACTC (SEQ ID NO:66) using a
10 fragment containing the murine neurturin gene as template
using (PCR parameters: 94°C for 30 sec; 60°C for 15 sec
and 68°C for 90 sec for 25 cycles with 5%= DMSO added
into the reaction). The amplified product was subcloned
into the Eco RV site of BSKS, the nucleotide sequence was
15 verified, and this fragment was then transferred to the
expression vector pET-30a (Novagen, Madison, WI) using an
Nde 1 site (5' end) and an Eco R1 site (3' end). The
pET-neurturin (pET-NTN) vector codes for an initiator
methionine in front of the first amino acid of the mature
20 mouse neurturin protein predicted from the N-terminal
amino acid sequence of neurturin purified from the CHO
cell conditioned media.

Example 7

25   This example illustrates the transient transfection
of NIH3T3 cells with the neurturin expression vector
pCMV-NTN-3-1 and that the product of the genomic sequence
in Example 5 is biologically active.

To demonstrate that the cloned neurturin cDNA was
30 sufficient to direct the synthesis of biologically active
neurturin we transiently introduced the pCMV-NTN-3-1
plasmid into NIH3T3 cells using the lipofectamine method
of transfection. NIH3T3 cells were plated at a density
of 400,000 cells per well (34.6 mm diameter) in 6 well
35 plates (Corning, Corning, NY) 24 hours before
transfection. DNA liposome complexes were prepared and

added to the cells according to the manufacturer's
protocol using 1.5 µg CMV-n urturin plasmid DNA (isolated
and purified using a Qiagen (Chatsworth, CA) tip-500
column according to manufacturer's protocol) and 10 µl
5   lipofectamine reagent (Gibco BRL, Gaithersburg, MD) in
1:1 DME/F12 medium containing 5 µg/ml insulin, 5 µg/ml
transferrin, and 5 ng/ml sodium selenite (Sigma, St.
Louis, MO). Five hours after the addition of DNA
liposome complexes in 1 ml medium per well, 1 ml DME
10   medium containing 20% calf serum was added to each well.
Twenty-four hours after the addition of DNA-liposome
complexes, the 2 ml medium above was replaced with 1 ml
DME medium containing 10% calf serum, 2 mM glutamine, 100
U/ml penicillin, 100 µ/ml streptomycin, and 25 ug/ml
15   heparin. The cells were incubated for an additional 24
hours before the conditioned medium was harvested,
centrifuged to remove cellular debris, and frozen.
        As a control, NIH3T3 cells were transfected as above
using 1.5 µg CMV-neo expression plasmid (containing no
20   cDNA insert) in place of the 1.5 µg CMV-neurturin
plasmid. Conditioned medium from NIH3T3 cells
transfected with either control plasmid or CMV-neurturin
plasmid was assayed by direct addition to the SCG culture
medium at the time of NGF deprivation. Addition of 0.25
25   ml conditioned medium from CMV-neurturin-transfected
cells promoted 70% survival of sympathetic neurons, and
>90% survival could be obtained with 0.45 ml of this
conditioned medium. No significant survival promoting
activity was detected in the conditioned medium of
30   control transfected NIH3T3 cells.

Example 8

        This example illustrates the preparation of Chinese
hamster ovary cells stably transformed with neurturin
35   cDNA.

94

DG44 cells, a Chinese hamster ovary cell derivative
that is deficient in dihydrofolate reductase (DHFR)
(Urlaub et al *Cell* 3:405-412, 1983 which is incorporated
by reference), were stably co-transfected with expression
5   plasmid (pCMV-NTN-3-1) and a DHFR expression plasmid
(HLD) (McArthur, and Stanners *J. Biol. Chem.*
266:6000-6005, 1991 which is incorporated by reference).

On day 1 DG44 cells were plated at 1x10^6 cells per
10 cm plate in Ham's F12 medium with 10% fetal calf serum
10 (FCS). This density must not be exceeded or cells will
overgrow before selection media is added on day 5.

On day 2 cells were transfected with a 9:1 ratio of
pCMV-NTN to DHFR expression plasmid using the calcium
phosphate method (10 ug DNA /10 cm plate) (Chen and
15 Okayama, *Mol Cell Biol* 7:2745-2752, 1987 which is
incorporated by reference).

On day 3 the transfected cells were washed with
Ham's F12 medium and fed Ham's F12 with 10% FCS.

On day 5 the cells were washed with MEM alpha medium
20 and fed selection medium, which is MEM alpha with 10% FCS
and 400 ug/ml G418. The cells were maintained in
selection media, feeding every 4 days. Colonies began to
appear approximately 14 days after transfection.
Colonies growing in selection media were then transferred
25 to a 24 well plate and trypsinized the next day to
disperse the cells. The cells were grown to confluence
in either 24 well or 6 well plates in order to screen the
cells for expression of recombinant protein. Expression
of neurturin was examined in 10 clonal lines and two high
30 expressing lines were detected using the SCG survival
assay. These clonal lines were expanded and expression
in these selected cell lines was amplified by selection
in 50 nM methotrexate (MTX). For selection in MTX, cells
were grown to 50% confluence in a 150 cm^2 flask in
35 selection medium. The medium was changed to MEM alpha
containing 50 nM MTX concentration (it was not necessary

- 95

to use G418 during MTX amplification). After placement
in 50 nM MTX, the majority of cells died and colonies of
resistant cells reappeared in 1-2 weeks. At this time,
the cells were trypsinized to disperse colonies and are
5 split when cells reach confluence. Cells eventually
reached the same growth rate as before. The selected
cells were screened for expression of recombinant
protein. A 2-3 fold increase in expression was observed
after selection in 50 nM MTX. Frozen stocks were kept
10 for cell lines obtained from the original selection and
the 50 nM MTX selection. Further selection could be
continued in increasing MTX until desired levels of
expression are obtained.

Using the above method, we isolated cells identified
15 as DG44CHO5-3(G418)(pCMV-NTN-3-1) and DG44CHO5-
3(50nMMTX)(pCMV-NTN-3-1). Cells from the DG44CHO5-
3(50nMMTX)(pCMV-NTN-3-1) strain expressed levels of
approximately 100 µg of biologically active protein per
liter of conditioned media determined by direct assay of
20 conditioned medium in SCG assay according to the methods
in example 1.

Example 9

This example illustrates the expression of neurturin
25 in various tissues.

A survey of neurturin and GDNF expression was
performed in rat embryonic tissues (E10, day 10 after
conception), neonatal tissues (P1, Postnatal Day 1), and
adult tissues (> 3 mos) using semi-quantitative RT/PCR
30 (Estus et al., *J Cell Biol* 127:1717-1727, 1994 which is
incorporated by reference). The RNA samples were
obtained from various tissues and PCR products were
detected either by autoradiography after incorporation of
α-^{32}P-dCTP in the PCR and electrophoresis on a
35 polyacrylamide gel (Figure 6) or by ethidium bromide
staining of DNA after electrophoresis on agarose gels

96

(Tables 3 and 4). The neurturin fragment of 101 base
pairs was obtained using the forward primer
CAGCGACGACGCGTGCGCAAAGAGCG (SEQ ID NO:67) and reverse
primer TAGCGGCTGTGTACGTCCAGGAAGGACACCTCGT (SEQ ID NO:68)

5 and the GDNF fragment of 194 base pairs was obtained
using the forward primer AAAAATCGGGGGTGYGTCTTA (SEQ ID
NO:69) and the reverse primer CATGCCTGGCCTACYTTGTCA (SEQ
ID NO:70).

       No neurturin or GDNF mRNA was detected at the
10 earliest embryonic age (embryonic day 10, E10) surveyed.

       In neonates (postnatal day 1, P1) both transcripts
were expressed in many tissues although neurturin tended
to show a greater expression in most tissues than did
GDNF. (see table 3).

15

Table 3.

	NEURTURIN	GDNF
Liver	+++	-
Blood	+++	+
Thymus	+	-
Brain	++	+
Sciatic nerve	-	+
Kidney	++	++
Spleen	++	+
Cerebellum	++	+
Heart	++	+
Bone	+	+

30   As shown in Table 3, differences in the tissue
distributions of neurturin and GDNF were noted. In
particular, no GDNF was detected in liver and thymus
where neurturin expression was detected and no neurturin
was detected in sciatic nerve where GDNF was detected.

35   Neurturin and GDNF mRNA were detected in many
tissues in the adult animal, but the tissue-specific

pattern of expression for these two genes was very
different. (table 4, Figure 5).

Table 4.

	NEURTURIN	GDNF
Liver	−	−
Blood	+	−
Thymus	+	++
Brain	+	−
Sciatic nerve	−	−
Kidney	++	+
Spleen	−	+
Cerebellum	−	−
Uterus	++	−
Bone marrow	++	−
Testis	++	++
Ovary	+	+
Placenta	+	−
Skeletal muscle	+	−
Spinal cord	+	−
Adrenal gland	++	++
Gut	+	++

As shown in table 4, neurturin was found to be
expressed in brain and spinal cord as well as in blood
and bone marrow where no GDNF was detected. The level of
expression of neurturin in brain and blood was, however,
less than that detected in neonatal tissue.

Neurturin was also highly expressed in freshly
isolated rat peritoneal mast cells, whereas GDNF showed
little or no expression.

Example 10

98

This example illustrates the preparation of antisera to neurturin by immunization of rabbits with a neurturin peptide.

The peptide sequence corresponding to amino acids
5 73-87 of the mature murine neurturin protein was synthesized and coupled to keyhole limpet hemocyanin (KLH) as described earlier (Harlow and Lane, *Antibodies: a laboratory manual*, 1988. Cold Spring Harbor Laboratory, New York, NY. p. 72-81 which is incorporated by
10 reference). The KLH-coupled peptide was submitted to Caltag, Inc. and each of two rabbits were immunized. Immunization was by subcutaneous injection at 7-10 sites. The first injection was with 150 µg KLH-coupled peptide which was resuspended in 0.5 ml saline and emulsified
15 with 0.5 ml complete Freund's adjuvant. Boost injections were begun 4 weeks after the initial injection and were performed once every 7 days as above for a total of 5 injections except that 100 µg of KLH-coupled peptide and incomplete Freund's adjuvant were used. Serum samples
20 were collected 1 week after the fifth boost.

A pooled volume of twenty ml of serum that had been collected from both rabbits one week after the 5th injection was purified. For purification, a peptide affinity column was prepared by coupling the above
25 peptide to cyanogen bromide activated Sepharose 4B according to the manufacturers protocol (Pharmacia Biotech). The serum was diluted 10 fold in 10 mM Tris pH 7.5 buffer and mixed by gentle rocking for 16 hours at 4°C with 0.5 ml of peptide agarose matrix containing 5 mg
30 of coupled peptide. The matrix was placed into a column, washed with 5 ml of 10 mM Tris pH 7.5, 150 mM NaCl, washed with 5 ml of 10 mM Tris pH 7.5 buffer containing 0.4 M NaCl and eluted with 5.5 ml of 100 mM glycine pH 2.5 buffer. One tenth volume of 1.0M Tris pH 8.0 buffer
35 was added to the eluate immediately after elution to

99

neutralize the pH. The glycine eluate was dialyzed overnight against 10 mM Tris pH 7.5, 150 mM NaCl.

The affinity-purified antibodies were used in a western blot to demonstrate specific recognition of
5 recombinant neurturin protein. Ten ml of conditioned medium collected from DG44CHO5-3(G418)(pCMV-NTN-3-1) cells was purified over SP Sepharose as described in Example 1 and the proteins electrophoresed on a reducing SDS-PAGE gel in the tricine buffer system (Schagger and
10 von Jagow *Analytical Biochemistry* 166:368-379, 1987). The proteins were electroblotted to a nitrocellulose membrane in 25 mM Tris, 192 mM glycine, 0.04% SDS, 17% methanol at 4°C for 16 hr. The membrane was incubated with the affinity-purified anti-neurturin peptide
15 antibodies and then with horseradish peroxidase-coupled sheep anti-rabbit IgG (Harlow and Lane, *supra*, p. 498-510). Bound antibodies were detected with enhanced chemiluminescence (ECL kit, Amersham, Buckinghamshire, England). The anti-neurturin antibodies recognized a
20 single, approximately 11.5 kD protein band in the conditioned medium of the DG44CHO5-3(G418)(pCMV-NTN-3-1) cells. Using these anti-neurturin antibodies, neurturin protein could be detected in 10 ml of conditioned medium from DG44CHO5-3(G418)(pCMV-NTN-3-1) cells but could not
25 be detected in 10 ml of medium conditioned with DG44 cells that had not been transformed with the neurturin expression vector.

Example 11
30   The following example illustrates the identification of additional members of the GDNF/neurturin/persephin gene subfamily.

The TGF-β superfamily currently contains over 25 different gene members (for review see Kingsley, *Genes*
35 *and Development* 8: 133-146, 1994 which is incorporated by reference). The individual family members display

100

varying degrees of homology with each other and several
subgroups within the superfamily can be defined by
phylogenetic analysis using the Clustal V program
(Higgins et al, *Comput Appl Biosci 8*: 189-191, 1992 which
5   is incorporated by reference) and by bootstrap analysis
of phylogenetic trees (Felsenstein, *Evolution 39*:783-791,
1985 which is incorporated by reference). Neurturin or
persephin is approximately 40% identical to GDNF but less
than 20% identical to any other member of the TGF-ß
10 superfamily. Several sequence regions in neurturin can
be identified (Figure 5) that are highly conserved within
the GDNF/neurturin/persephin subfamily but not within the
TGF-ß superfamily. These conserved regions are likely to
characterize a subfamily containing previously unisolated
15 genes, which can now be isolated using the conserved
sequence regions identified by the discovery and
sequencing of the neurturin and persephin genes. Regions
of high sequence conservation between neurturin,
persephin and GDNF allow the design of degenerate
20 oligonucleotides which can be used either as probes or
primers. Conserved-region amino acid sequences have been
identified herein to include Val-Xaa$_1$-Xaa$_2$-Leu-Gly-Leu-
Gly-Tyr where Xaa$_1$ is Ser, Thr or Ala and Xaa$_2$ is Glu or
Asp (SEQ ID NO:108); Glu-Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Arg-Tyr-Cys-
25 Xaa$_4$-Gly-Xaa$_5$-Cys in which Xaa$_1$ is Thr, Glu or lys, Xaa$_2$ is
Val, Leu or Ile, Xaa$_3$ is Leu or Ile, Xaa$_4$ is Ala or Ser,
and Xaa$_5$ is Ala or Ser, (SEQ ID NO:113); and Cys-Cys-Xaa$_1$-
Pro-Xaa$_2$-Xaa$_3$-Xaa$_4$-Xaa$_5$-Asp-Xaa$_6$-Xaa$_7$-Xaa$_8$-Phe-Leu-Asp-Xaa$_9$-
in which Xaa$_1$ is Arg or Gln, Xaa$_2$ is Thr or Val or Ile,
30 Xaa$_3$ is Ala or Ser, Xaa$_4$ is Tyr or Phe, Xaa$_5$ is Glu, Asp
or Ala, Xaa$_6$ is Glu, Asp or no amino acid, Xaa$_7$ is val or
leu, Xaa$_8$ is Ser or Thr, and Xaa$_9$ is Asp or Val (SEQ ID
NO:114). Nucleotide sequences containing a coding
sequence for the above conserved sequences or fragments
35 of the above conserved sequences can be used as probes.

101

Exemplary probe and primer sequences which can be designed from these regions are as follows.

Forward primers,

Primer A (M3119): 5'-GTNDGNGANYTGGGNYTGGGNTA (SEQ ID NO:115) 23 nt which codes for the amino acid sequence, Val-Xaa₁-Xaa₂-Leu-Gly-Leu-Gly-Tyr where Xaa₁ is Thr, Ser or Ala and Xaa₂ is Glu or Asp (SEQ ID NO:125);

Primer B (M3123): 5'-GANBTNWCNTTYYTNGANG (SEQ ID NO:116) 19 nt which codes for the amino acid sequence, Xaa₁-Xaa₂-Xaa₃-Phe-Leu-Xaa₄-Xaa₅ where Xaa₁ is Asp or Glu, Xaa₂ is Val or Leu, Xaa₃ is Thr or Ser, Xaa₄ is Asp or Glu, and Xaa₅ is Asp or Val (SEQ ID NO:126);

Primer C (M3126): 5'-GANBTNWCNTTYYTNGANGW (SEQ ID NO:117) 20 nt which codes for the amino acid sequence, Xaa₁-Xaa₂-Xaa₃-Phe-Leu-Xaa₄-Xaa₅ where Xaa₁ is Asp or Glu, Xaa₂ is Val or Leu, Xaa₃ is Thr or Ser, Xaa₄ is Asp or Glu, and Xaa₅ is Asp or Val (SEQ ID NO:126);

Primer D (M3121): 5'-TTYMGNTAYTGYDSNGGNDSNTG (SEQ ID NO:118) 23 nt which codes for the amino acid sequence, Phe-Arg-Tyr-Cys-Xaa₁-Gly-Xaa₂-Cys where Xaa₁ is Ser or Ala and Xaa₂ is Ser or Ala (SEQ ID NO:127);

Primer E (M3122): 5'-GTNDGNGANYTGGGNYTNGG (SEQ ID NO:119) 20 nt which codes for the amino acid sequence, Val-Xaa₁-Xaa₂-Leu-Gly-Leu-Gly where Xaa₁ is Thr, Ser or Ala and Xaa₂ is Asp or Glu (SEQ ID NO:128); and

102

Primer F (M3176): 5'-GTNDGNGANYTGGGNYTGGGNTT (SEQ ID NO:120) 23 nt which codes for the amino acid sequence, Val-Xaa$_1$-Xaa$_2$-Leu-Gly-Leu-Gly-Phe where Xaa$_1$ is Thr, Ser or Ala and Xaa$_2$ is Glu or Asp (SEQ ID NO:129).

Reverse primers,
Primer G (M3125): 5'-WCNTCNARRAANGWNAVNTC (SEQ ID NO:121) 20 nt whose reverse complementary sequence codes for the amino acid sequence, Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Leu-Xaa$_4$-Xaa$_5$ where Xaa$_1$ is Asp or Glu, Xaa$_2$ is Val or Leu, Xaa$_3$ is Thr or Ser, Xaa$_4$ is Asp or Glu, and Xaa$_5$ is Asp or Val (SEQ ID NO:126);

Primer H (M3124): 5'-WCNTCNARRAANGWNAVNT (SEQ ID NO:122) 19 nt whose reverse complementary sequence codes for the amino acid sequence, Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Leu-Xaa$_4$-Xaa$_5$ where Xaa$_1$ is Asp or Glu, Xaa$_2$ is Val or Leu, Xaa$_3$ is Thr or Ser, Xaa$_4$ is Asp or Glu, and Xaa$_5$ is Asp or Val (SEQ ID NO:126);

Primer I (M3120): 5'-CANSHNCCNSHRCARTANCKRAA (SEQ ID NO:123) 23 nt whose reverse complementary sequence codes for the amino acid sequence, Phe-Arg-Tyr-Cys-Xaa$_1$-Gly-Xaa$_2$-Cys where Xaa$_1$ is Ser or Ala and Xaa$_2$ is Ser or Ala (SEQ ID NO:127); and

Primer J (M3118): 5'-CANSHNCCNSHRCARTANCKRAANA (SEQ ID NO:124) 25 nt whose reverse complementary sequence codes for the amino acid sequence, Xaa$_1$-Phe-Arg-Tyr-Cys-Xaa$_2$-Gly-Xaa$_3$-Cys where Xaa$_1$ is Ile or Leu, Xaa$_2$ is Ser or Ala and Xaa$_3$ is Ser or Ala (SEQ ID NO:130).

103

In addition to the above, the following primers are based upon conserved regions in GDNF and neurturin (SEQ ID NOS:33-35).

Primer 1, GTNWSNGANYTNGGNYTNGGNTA (SEQ ID NO:42) which encodes the amino acid sequence, Val-Xaa$_1$-Xaa$_2$-Leu-Gly-Leu-Gly-Tyr where Xaa$_1$ is Ser or Thr and Xaa$_2$ is Glu or Asp (SEQ ID NO:33);

Primer 2, TTYMGNTAYTGYDSNGGNDSNTGYGANKCNGC (SEQ ID NO:43) which encodes amino acid sequence Phe-Arg-Tyr-Cys-Xaa$_1$-Gly-Xaa$_2$-Cys-Xaa$_3$-Xaa$_4$-Ala where Xaa$_1$ is Ala or Ser, Xaa$_2$ is Ala or Ser, Xaa$_3$ is Glu or Asp and Xaa$_4$ is Ser or Ala (SEQ ID NO:36);

Primer 3 reverse GCNGMNTCRCANSHNCCNSHRTANCKRAA (SEQ ID NO:44) whose reverse complementary sequence encodes amino acid sequence Phe-Arg-Tyr-Cys-Xaa$_1$-Gly-Xaa$_2$-Cys-Xaa$_3$-Xaa$_4$-Ala where Xaa$_1$ is Ala or Ser, Xaa$_2$ is Ala or Ser, Xaa$_3$ is Glu or Asp and Xaa$_4$ is Ser or Ala (SEQ ID NO:37);

Primer 4 reverse TCRTCNTCRWANGCNRYNGGNCKCARCA (SEQ ID NO:45) whose reverse complementary sequence encodes amino acid sequence Cys-Cys-Arg-Pro-Xaa$_1$-Ala-Xaa$_2$-Xaa$_3$-Asp-Xaa$_4$ where Xaa$_1$ is Ile or Thr or Val, Xaa$_2$ Try or Phe, Xaa$_3$ is Glu or Asp and Xaa$_4$ is Glu or Asp (SEQ ID NO:38);

Primer 5 reverse TCNARRAANSWNAVNTCRTCNTCRWANGC (SEQ ID NO:46) whose reverse complementary sequence encodes amino acid sequence Ala-Xaa$_1$-Xaa$_2$-Asp-Xaa$_3$-Xaa$_4$-Ser-Phe-Leu-Asp where Xaa$_1$ is Tyr or Phe, Xaa$_2$ Glu or Asp, Xaa$_3$ is Glu or Asp, and Xaa$_4$ is Val or Leu (SEQ ID NO:39);

Primer 6 GARRMNBTNHTNTTYMGNTAYTG (SEQ ID NO:47)
which encodes amino acid sequence Glu-Xaa$_1$-Xaa$_2$-
Xaa$_3$-Phe-Arg-Tyr-Cys where Xaa$_1$ is Glu or Thr, Xaa$_2$
is Leu or Val and Xaa$_3$ is Ile or Leu (SEQ ID
NO:40);

Primer 7 GARRMNBTNHTNTTYMGNTAYTGYDSNGGNDSNTGHGA
(SEQ ID NO:48) which encodes amino acid sequence
Glu-Xaa$_1$-Xaa$_2$-Xaa$_3$-Phe-Arg-Tyr-Cys-Xaa$_4$-Gly-Xaa$_5$-
Cys-Xaa$_6$ where Xaa$_1$ is Glu or Thr, Xaa$_2$ is Leu or
Val, Xaa$_3$ is Ile or Leu, Xaa$_4$ is Ser or Ala, Xaa$_5$
is Ser or Ala and Xaa$_6$ is Glu or Asp (SEQ ID
NO:41).

The above sequences can be used as probes for
screening libraries of genomic clones or as primers for
amplifying gene fragments from genomic DNA or libraries
of genomic clones or from reverse transcribed cDNA using
RNA templates from a variety of tissues. Genomic DNA or
libraries of genomic clones can be used as templates
because the neurturin, persephin and GDNF coding
sequences for the mature proteins are not interrupted by
introns.

A degenerate oligonucleotide can be synthesized as
a mixture of oligonucleotides containing all of the
possible nucleotide sequences which code for the
conserved amino acid sequence. To reduce the number of
different oligonucleotides in a degenerate mix, an
inosine or universal base (Loakes et al, Nucleic Acids
Res 22:4039-43, 1994) can be incorporated in the
synthesis at positions where all four nucleotides are
possible. The inosine or universal base forms base pairs
with each of the four normal DNA bases which are less
stabilizing than AT and GC base pairs but which are also
less destabilizing than mismatches between the normal
bases (i.e. AG, AC, TG, TC).

105

To isolate family members a primer above can be
end labeled with ³²P using T4 polynucleotide kinase and
hybridized to libraries of human genomic clones according
to standard procedures.

5   A preferred method for isolating family member
genes would be to use various combinations of the
degenerate primers above as primers in the polymerase
chain reaction using genomic DNA as a template. The
various combinations of primers can include sequential
10 PCR reactions utilizing nested primers or the use of a
forward primer paired with an oligo dT primer. In
addition, one of the degenerate primers can be used with
a vector primer, a single primer can be used in an
inverted PCR assay or PCR can be performed with one
15 degenerate primer and a random primer. As an example
using the above set of primers, primer 2 (SEQ ID NO:43)
can be used with primer 4 (SEQ ID NO:45) in PCR with 1 ug
of human genomic DNA and cycling parameters of 94°C for
30 sec, 50°C for 30 sec, and 72°C for 60 sec. The above
20 PCR conditions are exemplary only and one skilled in the
art will readily appreciate that a range of suitable
conditions and primer combinations could be used or
optimized such as different temperatures and varying salt
concentrations in the buffer medium and the like. It is
25 preferred that DMSO be added to the PCR reaction to a
final concentration of 5% inasmuch as this was found to
be necessary for amplification of this region of the
neurturin gene. The PCR reaction, when run on an agarose
gel, should contain products in the size range of 100-150
30 base pairs since a one amino acid gap is introduced in
the neurturin sequence and a five amino acid gap is
introduced in the persephin sequence when either sequence
is aligned with GDNF, and thus family member genes might
also contain a slightly variable spacing between the
35 conserved sequences of primers 2 and 4. The PCR products
in the range of 100-150 base pairs should contain

multiple amplified gene products including GDNF,
neurturin and persephin as well as previously unisolated
family members. To identify sequences of these products,
they can be gel purified and ligated into the Bluescript

5 plasmid (Stratagene), and then transformed into the
XL1-blue E. Coli host strain (Stratagene). Bacterial
colonies containing individual subclones can be picked
for isolation and plated on nitrocellulose filters in two
replicas. Each of the replicate filters can be screened

10 with an oligonucleotide probe for either unique GDNF or
unique neurturin or unique persephin sequence in the
amplified region. Subclones not hybridizing to either
GDNF or neurturin or persephin can be sequenced and if
found to encode previously unisolated family members, the

15 sequence can be used to isolate full length cDNA clones
and genomic clones as was done for neurturin (Example 5).
A similar method was used to isolate new gene members
(GDF-3 and GDF-9) of the TGF-ß superfamily based on
homology between previously identified genes (McPherron J

20 Biol Chem 268: 3444-3449, 1993 which is incorporated by
reference).

The inventors herein believe that the most
preferred way to isolate family member genes may be to
apply the above PCR procedure as a screening method to

25 isolate individual family member genomic clones from a
library. This is because there is only one exon for the
coding region of both mature neurturin and GDNF. If, for
example, the above PCR reaction with primers 2 and 4
generates products of the appropriate size using human

30 genomic DNA as template, the same reaction can be
performed using, as template, pools of genomic clones in
the P1 vector according to methods well known in the art,
for example that used for isolating neurturin human
genomic clones (Example 5). Pools containing the

35 neurturin gene in this library have previously been
identified and pesephin and GDNF-containing pools can be

107

readily identified by screening with GDNF specific
primers. Thus non-neurturin, non-persephin, non-GDNF
pools which generate a product of the correct size using
the degenerate primers will be readily recognized as
5　previously unisolated family members. The PCR products
generated from these pools can be sequenced directly
using the automated sequencer ~~and genomic~~ clones can be.
isolated by further subdivision and screening of the
pooled clones as a standard service offered by Genome
10　Systems, Inc.

Example 12

The following example illustrates the isolation
and identification of persephin utilizing the procedures
15　and primers described in Example 11.

The degenerate PCR strategy devised by the
inventors herein has now been successfully utilized to
identify a third factor, persephin, that is approximately
35-50% identical to both GDNF and neurturin. The
20　experimental approach was described above and is provided
in greater detail as follows. Primers corresponding to
the amino acid sequence Val-Xaa1-Xaa2-Leu-Gly-Leu-Gly-Tyr
where Xaa1 is Ser or Thr and Xaa2 is Glu or Asp (SEQ ID
NO:33) [M1996; 5'-GTNWSNGANYTNGGNYTNGGNTA (SEQ ID NO:42)]
25　and Phe-Arg-Tyr-Cys-Xaa1-Gly-Xaa2-Cys-Xaa3-Xaa4-Ala where
Xaa1 is Ala or Ser, Xaa2 is Ala or Ser, Xaa3 is Glu or
Asp and Xaa4 is Ser or Ala (SEQ ID NO:37) [M1999; 5'-
GCNGMNTCRCANSHNCCNSHRCARTANCKRAA (SEQ ID NO:44)] were
used to amplify a 77 nt fragment from rat genomic DNA
30　using Klentaq enzyme and buffer under the following
conditions:　94°C for 30 sec; 44°C for 30 sec; 72°C for
30 sec for 40 cycles. The resulting product was
subcloned into the Bluescript KS plasmid and sequenced.
All nucleotide sequencing was performed using fluorescent
35　dye terminator technology per manufacturer's instructions
on an Applied Biosystems automated sequencer Model #373

108

(Applied Biosystems, Fost r City, CA). Plasmid DNA for
sequencing was prepared using the Wizard Miniprep kit
(Promega Corp., Madison, WI) according to the
manufacturer's instructions.

5   The sequence of one of the amplified products
predicted amino acid sequence data internal to the PCR
primers that was different from that of GDNF or neurturin
but had more than 20% identity with GDNF and neurturin,
whereas the sequences of others we obtained corresponded
10 to GDNF or neurturin, as would be expected. The novel
sequence was thought to identify a new member of this
family which we named persephin.

The sequence of this fragment internal to the
primers was 5'-TGCCTCAGAGGAGAAGATTATC (SEQ ID NO:90).
15 This encodes the last nucleotide of the Tyr codon, and
then encodes the amino acids: Ala-Ser-Glu-Glu-Lys-Ile-
Ile (SEQ ID NO:91). This sequence was then aligned with
the rat sequences of GDNF and neurturin. This analysis
confirmed that persephin was unique.

20   LGLGYETKEELIFRYC GDNF (rat) (SEQ ID NO:92)
     LGLGYTSDETVLFRYC NTN (rat) (SEQ ID NO:93)
     LGLGYASEEKIIFRYC PSP (rat) (SEQ ID NO:94)

25   To obtain additional persephin sequence, primers
containing portions of the unique 22 nt of the amplified
fragment above were used in the rapid amplification of
cDNA ends (RACE) technique (Frohman, M.A. Methods in
Enzymology 218:340-356, 1993) using the Marathon RACE kit
30 (CLONTECH, Palo Alto, CA) per the manufacturer's
instructions, except that first strand cDNA synthesis was
carried out at 50°C using Superscript II reverse
transcriptase (Gibco-BRL). Briefly, a double stranded
adaptor oligonucleotide was ligated to the ends of double
35 stranded cDNA synthesized from postnatal day 1 rat brain
mRNA. Using nested forward persephin PCR primers,
(10135; 5'-AGTCGGGGTTGGGGTATGCCTCA, SEQ ID NO:95 and
M2026; 5'-TATGCCTCAGAGGAGAAGATTATCTT SEQ ID NO:96) in

combination with primers to the ligated adaptor supplied
in the kit (AP1, AP2), the 3' end of the persephin cDNA
was amplified by two successive PCR reactions (1st:
10135 and AP1, using 94°C for 30 sec, 60°C for 15 sec and
5 68°C for 2 min for 35 cycles; 2nd: M2026 and AP2 using
94°C for 30 sec, 60 for 15 sec and 68°C for 2 min for 21
cycles). An approximately 350 nt fragment was obtained
from this PCR reaction and this fragment was directly
sequenced using primer M2026. The sequence of this 3'
10 RACE product resulted in a partial rat persephin cDNA
sequence of approximately 350 nt (SEQ ID NO:97). The
predicted amino acid sequence of this cDNA was compared
to that of GDNF and neurturin, and found to be
approximately 40% homologous to each of these proteins.
15 Importantly, the characteristic spacing of the cyteine
residues in members of the TGF-β superfamily was present.
Furthermore, in addition to the region of similarity
encoded by the degenerate primers used to isolate
persephin, another region of high homology shared between
20 GDNF and neurturin, but absent in other members of the
TGF-β superfamily, was also present in persephin

```
GDNF   ACCRPVAFDDDLSFLDD   (aa 60-76)  (SEQ ID NO:98)
NTN    PCCRPTAYEDEVSFKDV   (aa 61-77)  (SEQ ID NO:99)
25  PSP    PCCQPTSYAD-VTFLDD   (aa 57-72)  (SEQ ID NO:100)
```

(Amino acid numbering uses the first Cys residue as amino
acid 1).

30   With the confirmation that persephin was indeed a
new member of the GDNF/neurturin subfamily, we isolated
murine genomic clones of persephin to obtain additional
sequence information. Primers (forward, M2026; 5'-
TATGCCTCAGAGGAGAAGATTATCTT, SEQ ID NO:96 and reverse,
35 M3028; 5'-TCATCAAGGAAGGTCACATCAGCATA, SEQ ID NO:101)
corresponding to rat cDNA sequence were used in a PCR
reaction (PCR parameters: 94°C for 30 sec, 55°C for 15
sec and 72°C for 30 sec for 35 cycles) to amplify a 155
nt fragment from mouse genomic DNA which was homologous

110

to rat persephin cDNA sequence. These primers were then
used to obtain murine pers phin genomic clones from a
mouse 129/Sv library in a P1 bacteriophage vector
(library screening service of Genome Systems, Inc., St.

5   Louis, MO).

Restriction fragments (3.4 kb Nco I and a 3.3 kb Bam
H1) from this P1 clone containing the persephin gene were
identified by hybridization with a 210 nt fragment
obtained by PCR using mouse genomic DNA with primers

10 (forward, M2026; SEQ ID NO:96 and reverse, M3159; 5'-
CCACCACAGCCACAAGCTGCGGSTGAGAGCTG, SEQ ID NO:102) and PCR
parameters: 94°C for 30 sec, 55°C for 15 sec and 72°C for
30 sec for 35 cycles. The Nco I and Bam Hi fragments
were sequenced and found to encode a stretch of amino

15 acids corresponding to that present in the rat persephin
RACE product, as well as being homologous to the mature
regions of both neurturin and GDNF (Figure 11).

When the amino acid sequences of murine GDNF,
neurturin and persephin are aligned using the first

20 cysteine as the starting point (which is done because
alterations in the cleavage sites between family members
creates variability in the segments upstream of the first
cysteine), persephin (91 amino acids) is somewhat smaller
than either neurturin (95 amino acids) or GDNF (94 amino

25 acids). The overall identity within this region is about
50% with neurturin and about 40% with GDNF (Figure 12).

Further nucleotide sequencing of the murine
persephin Nco I fragment revealed the nucleotide sequence
of the entire murine persephin gene (SEQ ID NO:131;

30 Figure 17). An open reading frame extends from the
sequence coding for an initiator methionine up to a stop
codon at positions 244-246. However, somewhere in this
sequence there is an apparent anomaly such that the
sequence encoding the RXXR cleavage site (nucleotides at

35 positions 257-268) and the sequence corresponding to the
mature persephin protein (positions 269-556) are not co-

111

linear with this open reading frame. Instead, a second
reading frame encodes the cleavage site and the mature
persephin. The two cogent reading frames are shown in
Figure 17.

5   Additional sequencing of the rat persephin has also
been performed. Rat genomic fragments were amplified by
PCR using Klentaq and rat genomic DNA as a template.
The forward primer #40266 (5'-AATCCCCAGGACAGGCAGGGAAT;
SEQ ID NO:137) corresponding to a region upstream of the
10   mouse persephin gene and a reverse primer M3156 (5'-
CGGTACCCAGATCTTCAGCCACCACAGCCACAAGC, SEQ ID NO:138)
corresponding to a region within the mature rat persephin
sequence were used with the following parameters (95°C
for 15 sec, 55°C for 15 sec, 68°C for 45 sec x 30
15   cycles). The amplified product was kinased with T4
polynucleotide kinase, the ends were blunted with E. coli
DNA polymerase I (Klenow fragment), and cloned into BSKS
plasmid.

       Nucleotide sequencing was performed to establish the
20   sequence of the entire rat persephin gene (SEQ ID NO:134;
Figure 18). An open reading frame was found to extend
from the sequence coding for an initiator methionine up
to a stop codon at positions 244-246 as was seen with
murine persephin. As was also seen with murine
25   persephin, an anomaly was found to occur between the
sequence encoding the initiator methionine and that
encoding the cleavage site for the mature rat persephin
such that two cogent reading frames exist as indicated in
Figure 18. Irrespective of this anomaly, mammalian cells
30   express persephin from either the murine or rat full
length genomic sequence as illustrated below.

Example 13

       This example illustrates the preparation of a
35   bacterial expression vector for murine persephin and its

112

introduction into an *E. Coli* for expression of recombinant mature pers phin.

The persephin polynucleotide encoding the mature murine persephin protein which begins 5 amino acids
5   upstream of the first framework Cys residue (SEQ ID NO:80) was cloned into the pET expression vector pET-30a at the Nde I and Bgl II sites. This persephin polynucleotide was generated by PCR using the murine persephin P1 genomic clone as a template. A forward
10   primer M3157 (5'-GGACTATCATATGGCCCACCACCACCACCACCACCACCACGACGACGACGACAAGGC CTTGGCTGGTTCATGCCGA, SEQ ID NO:139) encoding an Nde I site, 8 histidine residues, and an enterokinase site, and a reverse primer M3156 (5'-
15   CGGTACCCAGATCTTCAGCCACCACAGCCACAAGC, SEQ ID NO:138), which corresponds to the sequence encoding the last 6 amino acid residues of the mature persephin sequence, the stop codon and a Bgl II site, were used. The PCR reaction conditions were 95°C for 15 sec, 55°C for 15
20   sec, 68°C for 60 sec x 25 cycles. This PCR product was subcloned into the EcoRV site of BSKS plasmid and sequenced to verify that it contained no mutations. The persephin sequence was then excised from this vector using Nde I and Bgl II and cloned into the Nde I (5') and
25   Bgl II (3') sites of the bacterial expression vector pET30a (Novagen, Madison, WI). This expression vector would, therefore, produce the mature form of the persephin protein possessing an amino terminal tag consisting of 8 histidine residues followed directly by
30   an enterokinase site.

The plasmid was introduced into E.coli strain BL21 (DE3). To produce persephin, bacteria harboring this plasmid were grown for 16 hr, harvested, and lysed using 6M guanidine-HCl, 0.1 M NaH_2PO_4, 0.01 M Tris at pH 8.0,
35   and recombinant persephin protein was purified from these lysates via chromatography over a Ni-NTA resin (Qiagen).

113

The protein was eluted using 3 column volumes of Buffer E containing 8 M urea, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris, at pH 4.5. The persephin was then renatured by dialysis in renaturation buffer consisting of 0.1 M NaH$_2$PO$_4$, 0.01 M
5   Tris at pH 8.3, 0.15 M NaCl, 3 mM cysteine, 0.02% Tween-20, 10% glycerol and containing decreasing concentrations of urea beginning with 4 M for 16 hr, followed by 2 M for 16 hr, 1M for 72 hr, and 0.5 M for 16 hr. The persephin concentration was then determined using a Dot Metric
10 assay (Geno Technology, St. Louis, MO) and stored at 4°C.
        This bacterially produced recombinant persephin was used as an immunogen in rabbits to produce antibodies to mature persephin. All of the immunogen injections and blood drawing were performed at Cal Tag Inc. (Healdsburg,
15 CA). The anti-persephin antiserum was demonstrated to specifically recognize persephin, but not neurturin or GDNF, using protein blot analysis. This persephin-specific antiserum was then used to detect persephin in lysates prepared from transfected COS cells.
20
Example 14
        This example illustrates the preparation of mammalian expression vectors containing the murine or rat persephin genes and their incorporation into mammalian
25 cell lines for the production of mature persephin. To construct the murine plasmid, a P1 clone containing the murine persephin gene was used as a template in a PCR assay. Primers were designed such that the resulting polynucleotide would contain the persephin gene extending
30 from the initiator Methionine codon to the stop codon 3' to the mature persephin coding sequence (SEQ ID NO:131). The PCR reaction utilized a forward primer M3175 (5'-TGCTGTCACCATGGCTGCAGGAAGACTTCGGA, SEQ ID NO:140) and reverse primer M3156 (5'-
35 CGGTACCCAGATCTTCAGCCACCACAGCCACAAGC, SEQ ID NO:138). To construct the analogous rat plasmid, rat genomic DNA was

114

used as a template in a PCR assay. The PCR reaction
utilized a forward primer M3175 (5'-
TGCTGTCACCATGGCTGCAGGAAGACTTCGGA, SEQ ID NO:140) and
reverse primer M3156 (5'-

5   CGGTACCCAGATCTTCAGCCACCACAGCCACAAGC, SEQ ID NO:138).
Both PCR reactions were carried out using Klentaq and the
following parameters: 95°C for 15 sec; 55°C for 15 sec,
68°C for 45 sec x 25 cycles. The amplified products were
kinased with T4 polynucleotide kinase, the ends were

10 blunted with E. coli DNA polymerase I (Klenow fragment),
and cloned into BSKS plasmid. Nucleotide sequencing was
performed to verify that the correct clone was obtained.
The rat and murine persephin polynucleotides were excised
using Sma I and Hind III and each cloned into a Asp718

15 (blunted) and Hind III sites of the mammalian expression
vector pCB6.
     COS monkey cells were transfected with either the
rat or murine persephin expression vectors (16 µg per 5 x
10⁵ cells) or the non-recombinant vector (pCB6) itself

20 using the calcium phosphate precipitation method (Chen
and Okayama, Mol Cell Biol 7:2745-2752, 1987 which is
incorporated by reference). Forty eight hr later the
cells were lysed in IP buffer containing 50 mM Tris at pH
7.5, 300 mM NaCl, 1% Triton X-100, 1% deoxycholate, 10 mM

25 EDTA, 0.1% SDS, 5 µg/ml leupeptin, 7 µg/ml pepstatin, and
250 µM PMSF. The samples were loaded onto a 15% SDS-
polyacrylamide gel and the proteins were separated by
electrophoresis. The proteins were then transferred to
nitrocellulose by electroblotting. This nitrocellulose

30 membrane was incubated with anti-persephin antibodies to
detect the presence of persephin in the lysates.
     As is shown in Figure 19, lysates from cells
transfected with either the rat or murine persephin
expression vectors, but not the lysate from cells

35 transfected with pCB6, contain high amounts of persephin.
The size of the persephin detected was approximately 14

kD which is consistent with the size predicted for the processed, i.e. mature form of persephin. This demonstrates that both the murine and rat persephin genes are capable of directing the synthesis of a properly
5 processed persephin molecule.

Example 15

The following example illustrates the methods that can be used for isolation and identification of human
10 persephin.

The identification of murine and rat persephin sequences now allows us to identify and isolate the human persephin gene. Due to the high conservation between human and rodent GDNF (approximately 95% identity) and
15 between human and rodent neurturin (approximately 90% identity), it is believed that a similarly close relationship (i.e. greater than 85% identity) will be present between rodent and human persephin.

Different strategies can be used to obtain the human
20 persephin gene. In one preferred strategy, human genomic and cDNA libraries are screened by hybridization to the murine and/or the rat persephin sequences that have been identified herein (SEQ ID NOS:79-83) or portions of these sequences. These DNA sequences or probes are labeled as
25 described above, for example, with 32P-dCTP using either random priming or polynucleotide kinase. Hybridization conditions are described above and various stringency conditions of hybridization are used. Stringency of hybridization is determined by a number of factors during
30 hybridization and during the washing procedure, including temperature, ionic strength, length of time and concentration of formamide. These factors are outlined in, for example, Sambrook et al. *supra*.

In an alternate preferred strategy, primers
35 corresponding to portions of the rat or the murine persephin nucleotide sequence (or derivatives thereof)

are employed in a PCR reaction using either human genomic DNA or cDNA reverse transcribed from RNA isolated from human tissue. As an example, forward primer M2026 (SEQ ID NO:96) and reverse primer, M3028 (SEQ ID NO:101) can

5 be used in a PCR reaction using various conditions as described above and human DNA templates to amplify a human persephin fragment. Primers that amplify such a fragment as confirmed by nucleotide sequencing of the fragment are then used to obtain human persephin clones.

10 The clones are identified by virtue of their producing the same amplified fragment following PCR with the selected primers in a human genomic library in a P1 bacteriophage vector (library screening service of Genome Systems, Inc., St. Louis, MO).

15   After obtaining positive clones by either of the methods above, these human persephin clones are isolated and fragments are subcloned into Bluescript KS plasmids and sequenced. Nucleotide sequencing is performed using fluorescent dye terminator technology on an Applied

20 Biosystems automated sequencer Model #373 (Applied Biosystems, Foster City, CA) according to manufacturer's instructions. Plasmid DNA for sequencing is prepared using the Wizard Miniprep kit (Promega Corp., Madison, WI) according to manufacturer's instructions. Sequences

25 of these human fragments which are orthologous to the rat and murine persephin sequences are then identified and the full length nucleotide sequence of human persephin is established from the sequences of these fragments.

30   Example 16
        This example illustrates the preparation of chimeric or hybrid polypeptide molecules that contain portions derived from persephin (PSP) and portions derived from neurturin (NTN).

35   As closely related members of the TGFβ family, each of persephin and neurturin is predicted to have a very

117

similar overall structure, yet while neurturin promotes
the survival of sympathetic neurons, the closely related
persephin does not. Two chimeras were produced by
essentially replacing portions of persephin with

5   neurturin, with the crossover point located between the
two adjacent, highly conserved third and fourth cysteine
residues. The first chimera, named PSP/NTN (SEQ ID
NO:141, Figure 20), contains the first 63 residues of
mature murine persephin combined with residues 68 through

10 100 of mature murine neurturin (using E. coli preferred
codons). To construct this molecule, two PCR reactions
were performed: 1) using the forward primer M2012 (5'-
TAATACGACTCACTATAGGGGAA, SEQ ID NO:142) and reverse
primer M2188 (5'-

15 TCGTCTTCGTAAGCAGTCGGACGGCAGCAGGGTCGGCCATGGGCTCGAC, SEQ ID
NO:143) and the pET30a-murine persephin plasmid as
template (see Example 13); and 2) using the forward
primer M2190 (5'-TGCTGCCGTCCGACTGCTTACGAAGACGA, SEQ ID
NO:144) and reverse primer M2186 (5'-

20 GTTATGCTAGTTATTGCTCAGCGGT, SEQ ID NO:145) and the pET30a-
murine (E.coli preferred codons) neurturin plasmid as
template (see Example 6). Both PCR reactions were
carried out using the following parameters: 94°C for 30
sec, 55°C for 30 sec, 72°C for 30 sec x 25 cycles. The

25 products of these two PCR reactions were gel purified,
mixed together, and a PCR reaction was performed under
the following conditions: 94°C for 30 sec, 60°C for 20
min, 68°C for 5 min. After 8 cycles, an aliquot of this
reaction was used as template in a third PCR reaction

30 using the forward primer M2012 and reverse primer M2186
under the following conditions: 94°C for 30 sec, 55°C for
30 sec, 72°C for 30 sec x 25 cycles. The resulting
product was kinased with T4 polynucleotide kinase, the
ends were blunted with E. coli DNA polymerase I (Klenow

35 fragment), and cloned into BSKS plasmid. Nucleotide
sequencing was performed to verify that the correct clone

118

was obtained. The PSP/NTN fragment was excised using Nde I and Bam H1 and cloned into the corr sponding sites of the bacterial expression vector pET30a.

The second chimera, named NTN/PSP (SEQ ID NO:146,
5 Figure 20), encodes the converse molecule. It contains the first 67 residues of mature murine neurturin (using E.coli preferred codons) combined with residues 64 to 96 of mature murine persephin. To construct this molecule, we performed two PCR reactions: 1) using the forward
10 primer M2012 and reverse primer M2183 (5'-CACATCAGCATAGCTGGTGGGCTGGCAGCACGGGTGAGCACGAGCACGTT, SEQ ID NO:147) and the pET30a-murine (E.coli preferred codons) neurturin plasmid as template; and 2) using the forward primer M2187 (5'-TGCTGCCAGCCCACCAGCTATGCTG, SEQ
15 ID NO:148) and reverse primer M2186 (5'-GTTATGCTAGTTATTGCTCAGCGGT, SEQ ID NO:145) and the pET30a-murine persephin plasmid as template. Both PCR reactions were carried out using the following parameters: 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec x 25 cycles.
20 The products of these two PCR reactions were used to construct the final NTN/PSP pET30a plasmid as detailed above for PSP/NTN except that Bgl II was used instead of Bam H1. These chimeric proteins were produced in E.coli and purified by Ni-NTA chromatography as described above
25 (Example 13).

The purified proteins were assayed for their ability to promote survival in the SCG sympathetic neuron assay. The NTN/PSP protein did not promote survival, whereas the PSP/NTN protein promoted the survival of sympathetic
30 neurons similar to that observed for neurturin itself. These results indicate that neurturin residues lying downstream of the 2 adjacent, highly conserved cysteine residues are critical for activity in promoting survival in SCG sympathetic neurons. In contrast, the
35 corresponding residues of persephin are not sufficient for promoting survival in sympathetic neurons.

119

Example 17

This example illustrates the neuronal survival promoting activity of persephin in mesencephalic cells.

The profile of survival promoting activity of
5   persephin is different from that of neurturin and GDNF.
In contrast to the survival promoting activity produced
by neurturin and GDNF in sympathetic and sensory neurons,
persephin showed no survival promoting activity in these
tissues. We further evaluated the neuronal survival
10 promoting activity of persephin in mesencephalic cells.

Timed-pregnant Sprague-Dawley rats were purchased
from Harlan Sprague-Dawley. The mesencephalon was taken
from rats measuring 1.2 to 1.4 cm in length and time
dated to be embryonic day 14. The cranium was removed
15 and the entire mesencephalon was placed in cold L15. The
pooled mesencephalic tissue was resuspended in a serum-
free medium consisting of DME/Hams F12 (#11330-032, Life
Technologies) 1 mg/ml BSA, Fraction V (A-6793, Sigma
Chemical Co.,), 5 µM Insulin (I-5500, Sigma), 10 nM
20 progesterone (P0130, Sigma), 100 µM putrescine, (p7505,
Sigma), 30 nM Selenium (S07150, Pflatz & Bauer), 10 ng/ml
rat transferrin (012-000-050, Jackson Chrompure), 100
U/ml penicillin, and 100 U/ml of streptomycin. The
pooled mesencephalic tissues were triturated
25 approximately 80 times using a bent-tip pipette and the
cells were plated in a 24-well dish (CoStar) at a density
of 15,000 cells in a 100-µl drop. The dishes were coated
with 125 ng/ml poly-d-lysine (p-7280, Sigma) and 25 ng/ml
laminin (#40232, Collaborative Biomedical Products).
30 These dissociated cells were allowed to attach for 2
hours at 37°C in 5% CO_2 and then fed with another 500 µl
of the above serum-free medium with or without
approximately 100 ng/ml of recombinant Persephin. These
cells were photographed after 3 days of culture.
35   Inspection of the cells over the course of 3 days in
culture, showed a gradual decrease in cell number. In

120

the absence of any growth factor, almost all of the cells were dead (Figure 21A). In the presence of persephin, a large increase in mesencephalic neuronal cell survival was evident (Figure 21B).

5

Example 18

This example illustrates the expression of persephin in various tissues.

A survey of persephin expression was performed in
10 adult mouse tissues using semi-quantitative RT/PCR (see Example 9). Poly A RNA was isolated from brain, cerebellum, kidney, lung, heart, ovary, sciatic nerve, dorsal root ganglia, blood and spleen. This was then reverse transcribed to produce cDNA (see Kotzbauer et al.
15 *Nature* 384:467-470, 1996 which is incorporated by reference). The PCR primers used were as follows: forward primer: 5'-CCTCGGAGGAGAAGGTCATCTTC (SEQ ID NO:149) and reverse primer: 5'TCATCAAGGAAGGTCACATCAGCATA (SEQ ID NO:101). PCR was done for 26 cycles with an
20 annealing temperature of 60°C. To control for the presence of genomic DNA, RNA samples which were not reverse transcribed were used for PCR (for example, the tissue control shown in figure 22 is labeled "Kidney no RT"). All the samples were found to be without genomic
25 DNA contamination.

As shown in Figure 22, a band of the correct size (160 bp) was seen in the kidney sample. At higher cycle numbers a persephin band was also seen in brain. Thus, the distribution of expression of persephin in various
30 mouse tissues differs from that of neurturin in rat (Example 8).

Deposit of Strain. The following strain is on deposit under the terms of the Budapest Treaty, with the American Type Culture Collection, 12301 Parklawn Drive, Rockville,
35 MD. The accession number indicated was assigned after successful viability testing, and the requisite fees were

121

paid. Access to said cultures will be available during
pendency of the patent application to one determined by
the Commissioner to be entitled thereto under 37 CFR 1.14
and 35 USC 122. All restriction on availability of said
5 cultures to the public will be irrevocably removed upon
the granting of a patent based upon the application.
Moreover, the designated deposits will be maintained for
a period of thirty (30) years from the date of deposit,
or for five (5) years after the last request for the
10 deposit, or for the enforceable life of the U.S. patent,
whichever is longer. Should a culture become nonviable
or be inadvertently destroyed, or, in the case of
plasmid-containing strains, lose its plasmid, it will be
replaced with a viable culture. The deposited materials
15 mentioned herein are intended for convenience only, and
are not required to practice the present invention in
view of the description herein, and in addition, these
materials are incorporated herein by reference.

Strain	Deposit Date	ATCC No.
DG44CHO-pHSP-NGFI-B	August 25, 1995	CRL 11977

20

25

In view of the above, it will be seen that the
several advantages of the invention are achieved and
other advantageous results attained.
30   As various changes could be made in the above
methods and compositions without departing from the scope
of the invention, it is intended that all matter
contained in the above description and shown in the
accompanying drawings shall be interpreted as
35 illustrative and not in a limiting sense.

122

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: JOHNSON JR., EUGENE M.
                  MILBRANDT, JEFFREY D.
                  KOTZBAUER, PAUL T.
                  LAMPE, PATRICIA A.

    (ii) TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS

   (iii) NUMBER OF SEQUENCES: 176

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: HOWELL & HAFERKAMP, L.C.
        (B) STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
        (C) CITY: ST. LOUIS
        (D) STATE: MISSOURI
        (E) COUNTRY: US
        (F) ZIP: 63105-1817

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

   (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: HOLLAND, DONALD R.
        (B) REGISTRATION NUMBER: 35,197
        (C) REFERENCE/DOCKET NUMBER: 976163

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (314) 727-5188
        (B) TELEFAX: (314) 727-6092

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Ala Arg Leu Gly Ala Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg
1               5                   10                  15

Val Ser Glu Leu Gly Leu Gly Tyr Ala Ser Asp Glu Thr Val Leu Phe
            20                  25                  30

Arg Tyr Cys Ala Gly Ala Cys Glu Ala Ala Ala Arg Val Tyr Asp Leu
        35                  40                  45

Gly Leu Arg Arg Leu Arg Gln Arg Arg Arg Leu Arg Arg Glu Arg Val
    50                  55                  60
```

123

Arg Ala Gln Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser
65   70   75   80

Phe Leu Asp Ala His Ser Arg Tyr His Thr Val His Glu Leu Ser Ala
             85   90   95

Arg Glu Cys Ala Cys Val
            100

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 100 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

    Pro Gly Ala Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser
    1   5   10   15

    Glu Leu Gly Leu Gly Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr
             20   25   30

    Cys Ala Gly Ala Cys Glu Ala Ala Ile Arg Ile Tyr Asp Leu Gly Leu
             35   40   45

    Arg Arg Leu Arg Gln Arg Arg Val Arg Arg Glu Arg Ala Arg Ala
        50   55   60

    His Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu
    65   70   75   80

    Asp Val His Ser Arg Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu
             85   90   95

    Cys Ala Cys Val
            100

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "ANY AMINO ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

    Ser Gly Ala Arg Pro Xaa Gly Leu Arg Glu Leu Glu Val Ser Val Ser
    1   5   10   15

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1
        (D) OTHER INFORMATION: /note= "ANY AMINO ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "SERINE OR CYSTEINE"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

    Xaa Cys Ala Gly Ala Xaa Glu Ala Ala Val
    1   5   10

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1
        (D) OTHER INFORMATION: /note= "ANY AMINO ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "ANY AMINO ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 17
        (D) OTHER INFORMATION: /note= "GLUTAMINE OR GLUTAMIC ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

    Xaa Xaa Val Glu Ala Lys Pro Cys Cys Gly Pro Thr Ala Tyr Glu Asp
    1   5   10   15

    Xaa Val Ser Phe Leu Ser Val
                20

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

125

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Tyr His Thr Leu Gln Glu Leu Ser Ala Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 197 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Gln Arg Trp Lys Ala Ala Ala Leu Ala Ser Val Leu Cys Ser Ser
1   5   10   15

Val Leu Ser Ile Trp Met Cys Arg Glu Gly Leu Leu Leu Ser His Arg
            20   25   30

Leu Gly Pro Ala Leu Val Pro Leu His Arg Leu Pro Arg Thr Leu Asp
        35   40   45

Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala
    50   55   60

Pro Asp Ala Met Glu Leu Arg Glu Leu Thr Pro Trp Ala Gly Arg Pro
65   70   75   80

Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Ala Arg Ala
                85   90   95

Arg Leu Gly Ala Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg Val
            100   105   110

Ser Glu Leu Gly Leu Gly Tyr Ala Ser Asp Glu Thr Val Leu Phe Arg
        115   120   125

Tyr Cys Ala Gly Ala Cys Glu Ala Ala Ala Arg Val Tyr Asp Leu Gly
    130   135   140

Leu Arg Arg Leu Arg Gln Arg Arg Arg Leu Arg Arg Glu Arg Val Arg
145   150   155   160

Ala Gln Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe
            165   170   175

Leu Asp Ala His Ser Arg Tyr His Thr Val His Glu Leu Ser Ala Arg
        180   185   190

Glu Cys Ala Cys Val
        195

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 195 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

126

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Arg Arg Trp Lys Ala Ala Ala Leu Val Ser Leu Ile Cys Ser Ser
1               5                   10                  15

Leu Leu Ser Val Trp Met Cys Gln Glu Gly Leu Leu Leu Gly His Arg
            20                  25                  30

Leu Gly Pro Ala Leu Ala Pro Leu Arg Arg Pro Pro Arg Thr Leu Asp
            35                  40                  45

Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala
    50                  55                  60

Pro Asp Ala Val Glu Leu Arg Glu Leu Ser Pro Trp Ala Ala Arg Ile
65                  70                  75                  80

Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Ala Arg Pro
            85                  90                  95

Gly Ala Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu
            100                 105                 110

Leu Gly Leu Gly Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys
            115                 120                 125

Ala Gly Ala Cys Glu Ala Ala Ile Arg Ile Tyr Asp Leu Gly Leu Arg
    130                 135                 140

Arg Leu Arg Gln Arg Arg Arg Val Arg Arg Glu Arg Ala Arg Ala His
145                 150                 155                 160

Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp
            165                 170                 175

Val His Ser Arg Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu Cys
            180                 185                 190

Ala Cys Val
            195
```

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 306 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
GCGCGGTTGG GGGCGCGGCC TTGCGGGCTG CGCGAGCTGG AGGTGCGCGT GAGCGAGCTG      60

GGCCTGGGCT ACGCGTCCGA CGAGACGGTG CTGTTCCGCT ACTGCGCAGG CGCCTGCGAG     120

GCTGCCGCGC GCGTCTACGA CCTCGGGCTG CGACGACTGC GCCAGCGGCG GCGCCTGCGG     180

CGGGAGCGGG TGCGCGCGCA GCCCTGCTGC CGCCCGACGG CCTACGAGGA CGAGGTGTCC     240

TTCCTGGACG CGCACAGCCG CTACCACACG GTGCACGAGC TGTCGGCGCG CGAGTGCGCC     300

TGCGTG                                                                306
```

127

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
CCGGGGGCTC GGCCTTGTGG GCTGCGCGAG CTCGAGGTGC GCGTGAGCGA GCTGGGCCTG        60

GGCTACACGT CGGATGAGAC CGTGCTGTTC CGCTACTGCG CAGGCGCGTG CGAGGCGGCC       120

ATCCGCATCT ACGACCTGGG CCTTCGGCGC CTGCGCCAGC GGAGGCGCGT GCGCAGAGAG       180

CGGGCGCGGG CGCACCCGTG TTGTCGCCCG ACGGCCTATG AGGACGAGGT GTCCTTCCTG       240

GACGTGCACA GCCGCTACCA CACGCTGCAA GAGCTGTCGG CGCGGGAGTG CGCGTGCGTG       300
```

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 591 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
ATGCAGCGCT GGAAGGCGGC GGCCTTGGCC TCAGTGCTCT GCAGCTCCGT GCTGTCCATC        60

TGGATGTGTC GAGAGGGCCT GCTTCTCAGC CACCGCCTCG GACCTGCGCT GGTCCCCCTG       120

CACCGCCTGC CTCGAACCCT GGACGCCCGG ATTGCCCGCC TGGCCCAGTA CCGTGCACTC       180

CTGCAGGGGG CCCCGGATGC GATGGAGCTG CGCGAGCTGA CGCCCTGGGC TGGGCGGCCC       240

CCAGGTCCGC GCCGTCGGGC GGGGCCCCGG CGGCGGCGCG CGCGTGCGCG GTTGGGGGCG       300

CGGCCTTGCG GGCTGCGCGA GCTGGAGGTG CGCGTGAGCG AGCTGGGCCT GGGCTACGCG       360

TCCGACGAGA CGGTGCTGTT CCGCTACTGC GCAGGCGCCT GCGAGGCTGC CGCGCGCGTC       420

TACGACCTCG GGCTGCGACG ACTGCGCCAG CGGCGGCGCC TGCGGCGGGA GCGGGTGCGC       480

GCGCAGCCCT GCTGCCGCCC GACGGCCTAC GAGGACGAGG TGTCCTTCCT GGACGCGCAC       540

AGCCGCTACC ACACGGTGCA CGAGCTGTCG GCGCGCGAGT GCGCCTGCGT G               591
```

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 585 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATGAGGCGCT GGAAGGCAGC GGCCCTGGTG TCGCTCATCT GCAGCTCCCT GCTATCTGTC   60

TGGATGTGCC AGGAGGGTCT GCTCTTGGGC CACCGCCTGG GACCCGCGCT TGCCCCGCTA   120

CGACGCCCTC CACGCACCCT GGACGCCCGC ATCGCCCGCC TGGCCCAGTA TCGCGCTCTG   180

CTCCAGGGCG CCCCCGACGC GGTGGAGCTT CGAGAACTTT CTCCCTGGGC TGCCCGCATC   240

CCGGGACCGC GCCGTCGAGC GGGTCCCCGG CGTCGGCGGG CGCGGCCGGG GGCTCGGCCT   300

TGTGGGCTGC GCGAGCTCGA GGTGCGCGTG AGCGAGCTGG GCCTGGGCTA CACGTCGGAT   360

GAGACCGTGC TGTTCCGCTA CTGCGCAGGC GCGTGCGAGG CGGCCATCCG CATCTACGAC   420

CTGGGCCTTC GGCGCCTGCG CCAGCGGAGG CGCGTGCGCA GAGAGCGGGC GCGGGCGCAC   480

CCGTGTTGTC GCCCGACGGC CTATGAGGAC GAGGTGTCCT TCCTGGACGT GCACAGCCGC   540

TACCACACGC TGCAAGAGCT GTCGGCGCGG GAGTGCGCGT GCGTG   585

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 348 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGAGGGAGAG CGCGCGGTGG TTTCGTCCGT GTGCCCCGCG CCCGGCGCTC CTCGCGTGGC   60

CCCGCGTCCT GAGCGCGCTC CAGCCTCCCA CGCGCGCCAC CCCGGGGTTC ACTGAGCCCG   120

GCGAGCCCGG GGAAGACAGA GAAAGAGAGG CCAGGGGGGG AACCCCATGG CCCGGCCCGT   180

GTCCCGCACC CTGTGCGGTG GCCTCCTCCG GCACGGGGTC CCCGGGTCGC CTCCGGTCCC   240

CGCGATCCGG ATGGCGCACG CAGTGGCTGG GGCCGGGCCG GGCTCGGGTG GTCGGAGGAG   300

TCACCACTGA CCGGGTCATC TGGAGCCCGT GGCAGGCCGA GGCCCAGG   348

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 87 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGCTACCTCA CGCCCCCCGA CCTGCGAAAG GGCCCTCCCT GCCGACCCTC GCTGAGAACT   60

GACTTCACAT AAAGTGTGGG AACTCCC   87

(2) INFORMATION FOR SEQ ID NO:15:

129

```
     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 19 amino acids
         (B) TYPE: amino acid
         (C) STRANDEDNESS:
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

    Met Gln Arg Trp Lys Ala Ala Ala Leu Ala Ser Val Leu Cys Ser Ser
    1               5                   10                  15

    Val Leu Ser


(2) INFORMATION FOR SEQ ID NO:16:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 19 amino acids
         (B) TYPE: amino acid
         (C) STRANDEDNESS:
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

    Met Arg Arg Trp Lys Ala Ala Ala Leu Val Ser Leu Ile Cys Ser Ser
    1               5                   10                  15

    Leu Leu Ser


(2) INFORMATION FOR SEQ ID NO:17:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 57 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATGCAGCGCT GGAAGGCGGC GGCCTTGGCC TCAGTGCTCT GCAGCTCCGT GCTGTCC          57

(2) INFORMATION FOR SEQ ID NO:18:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 57 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGAGGCGCT GGAAGGCAGC GGCCCTGGTG TCGCTCATCT GCAGCTCCCT GCTATCT          57
```

130

(2) INFORMATION FOR SEQ ID NO:19:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 76 amino acids
             (B) TYPE: amino acid
             (C) STRANDEDNESS:
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

        Ile Trp Met Cys Arg Glu Gly Leu Leu Leu Ser His Arg Leu Gly Pro
        1   5   10   15

        Ala Leu Val Pro Leu His Arg Leu Pro Arg Thr Leu Asp Ala Arg Ile
                    20   25   30

        Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala Pro Asp Ala
                35   40   45

        Met Glu Leu Arg Glu Leu Thr Pro Trp Ala Gly Arg Pro Pro Gly Pro
            50   55   60

        Arg Arg Arg Ala Gly Pro Arg Arg Arg Arg Ala Arg
        65   70   75

(2) INFORMATION FOR SEQ ID NO:20:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 228 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATCTGGATGT GTCGAGAGGG CCTGCTTCTC AGCCACCGCC TCGGACCTGC GCTGGTCCCC   60

CTGCACCGCC TGCCTCGAAC CCTGGACGCC CGGATTGCCC GCCTGGCCCA GTACCGTGCA   120

CTCCTGCAGG GGGCCCCGGA TGCGATGGAG CTGCGCGAGC TGACGCCCTG GGCTGGGCGG   180

CCCCCAGGTC CGCGCCGTCG GGCGGGGCCC CGGCGGCGGC GCGCGCGT   228

(2) INFORMATION FOR SEQ ID NO:21:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 228 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTCTGGATGT GCCAGGAGGG TCTGCTCTTG GGCCACCGCC TGGGACCCGC GCTTGCCCCG   60

CTACGACGCC CTCCACGCAC CCTGGACGCC CGCATCGCCC GCCTGGCCCA GTATCGCGCT   120

131

CTGCTCCAGG GCGCCCCCGA CGCGGTGGAG CTTCGAGAAC TTTCTCCCTG GGCTGCCCGC   180

ATCCCGGGAC CGCGCCGTCG AGCGGGTCCC CGGCGTCGGC GGGCGCGG   228

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 76 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Val Trp Met Cys Gln Glu Gly Leu Leu Leu Gly His Arg Leu Gly Pro
1   5   10   15

Ala Leu Ala Pro Leu Arg Arg Pro Pro Arg Thr Leu Asp Ala Arg Ile
              20   25   30

Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala Pro Asp Ala
          35   40   45

Val Glu Leu Arg Glu Leu Ser Pro Trp Ala Ala Arg Ile Pro Gly Pro
      50   55   60

Arg Arg Arg Ala Gly Pro Arg Arg Arg Arg Ala Arg
65   70   75

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 95 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Gln Arg Trp Lys Ala Ala Ala Leu Ala Ser Val Leu Cys Ser Ser
1   5   10   15

Val Leu Ser Ile Trp Met Cys Arg Glu Gly Leu Leu Leu Ser His Arg
              20   25   30

Leu Gly Pro Ala Leu Val Pro Leu His Arg Leu Pro Arg Thr Leu Asp
          35   40   45

Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala
      50   55   60

Pro Asp Ala Met Glu Leu Arg Glu Leu Thr Pro Trp Ala Gly Arg Pro
65   70   75   80

Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Arg Ala Arg
              85   90   95

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 95 amino acids

132

      (B) TYPE: amino acid
      (C) STRANDEDNESS:
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
Met Arg Arg Trp Lys Ala Ala Ala Leu Val Ser Leu Ile Cys Ser Ser
1               5                   10                  15

Leu Leu Ser Val Trp Met Cys Gln Glu Gly Leu Leu Leu Gly His Arg
            20                  25                  30

Leu Gly Pro Ala Leu Ala Pro Leu Arg Arg Pro Pro Arg Thr Leu Asp
            35                  40                  45

Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu Leu Gln Gly Ala
        50                  55                  60

Pro Asp Ala Val Glu Leu Arg Glu Leu Ser Pro Trp Ala Ala Arg Ile
65                  70                  75                  80

Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Ala Arg
                85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 285 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
ATGCAGCGCT GGAAGGCGGC GGCCTTGGCC TCAGTGCTCT GCAGCTCCGT GCTGTCCATC          60

TGGATGTGTC GAGAGGGCCT GCTTCTCAGC CACCGCCTCG GACCTGCGCT GGTCCCCCTG         120

CACCGCCTGC CTCGAACCCT GGACGCCCGG ATTGCCCGCC TGGCCCAGTA CCGTGCACTC         180

CTGCAGGGGG CCCCGGATGC GATGGAGCTG CGCGAGCTGA CGCCCTGGGC TGGGCGGCCC         240

CCAGGTCCGC GCCGTCGGGC GGGGCCCCGG CGGCGGCGCG CGCGT                         285
```

(2) INFORMATION FOR SEQ ID NO:26:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 285 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
ATGAGGCGCT GGAAGGCAGC GGCCCTGGTG TCGCTCATCT GCAGCTCCCT GCTATCTGTC          60

TGGATGTGCC AGGAGGGTCT GCTCTTGGGC CACCGCCTGG GACCCGCGCT TGCCCCGCTA         120
```

133

CGACGCCCTC CACGCACCCT GGACGCCCGC ATCGCCCGCC TGGCCCAGTA TCGCGGCTCTG   180

CTCCAGGGCG CCCCCGACGC GGTGGAGCTT CGAGAACTTT CTCCCTGGGC TGCCCGCATC   240

CCGGGACCGC GCCGTCGAGC GGGTCCCCGG CGTCGGCGGG CGCGG   285

(2) INFORMATION FOR SEQ ID NO:27:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 169 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGCAGCGCT GGAAGGCGGC GGCCTTGGCC TCAGTGCTCT GCAGCTCCGT GCTGTCCATC   60

TGGATGTGTC GAGAGGGCCT GCTTCTCAGC CACCGCCTCG GACCTGCGCT GGTCCCCCTG   120

CACCGCCTGC CTCGAACCCT GGACGCCCGG ATTGCCCGCC TGGCCCAGT   169

(2) INFORMATION FOR SEQ ID NO:28:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 425 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ACCGTGCACT CCTGCAGGGG GCCCCGGATG CGATGGAGCT GCGCGAGCTG ACGCCCTGGG   60

CTGGGCGGCC CCCAGGTCCG CGCCGTCGGG CGGGGCCCCG GCGGCGGCGC GCGCGTGCGC   120

GGTTGGGGGC GCGGCCTTGC GGGCTGCGCG AGCTGGAGGT GCGCGTGAGC GAGCTGGGCC   180

TGGGCTACGC GTCCGACGAG ACGGTGCTGT TCCGCTACTG CGCAGGCGCC TGCGAGGCTG   240

CCGCGCGCGT CTACGACCTC GGGCTGCGAC GACTGCGCCA GCGGCGGCGC CTGCGGCGGG   300

AGCGGGTGCG CGCGCAGCCC TGCTGCCGCC CGACGGCCTA CGAGGACGAG GTGTCCTTCC   360

TGGACGCGCA CAGCCGCTAC CACACGGTGC ACGAGCTGTC GGCGCGCGAG TGCGCCTGCG   420

TGTGA   425

(2) INFORMATION FOR SEQ ID NO:29:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 169 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

134

ATGAGGCGCT GGAAGGCAGC GGCCCTGGTG TCGCTCATCT GCAGCTCCCT GCTATCTGTC   60

TGGATGTGCC AGGAGGGTCT GCTCTTGGGC CACCGCCTGG GACCCGCGCT TGCCCCGCTA   120

CGACGCCCTC CACGCACCCT GGACGCCCGC ATCGCCCGCC TGGCCCAGT   169

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 419 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATCGCGCTCT GCTCCAGGGC GCCCCCGACG CGGTGGAGCT TCGAGAACTT TCTCCCTGGG   60

CTGCCCGCAT CCCGGGACCG CGCCGTCGAG CGGGTCCCCG GCGTCGGCGG GCGCGGCCGG   120

GGGCTCGGCC TTGTGGGCTG CGCGAGCTCG AGGTGCGCGT GAGCGAGCTG GGCCTGGGCT   180

ACACGTCGGA TGAGACCGTG CTGTTCCGCT ACTGCGCAGG CGCGTGCGAG GCGGCCATCC   240

GCATCTACGA CCTGGGCCTT CGGCGCCTGC GCCAGCGGAG GCGCGTGCGC AGAGAGCGGG   300

CGCGGGCGCA CCCGTGTTGT CGCCCGACGG CCTATGAGGA CGAGGTGTCC TTCCTGGACG   360

TGCACAGCCG CTACCACACG CTGCAAGAGC TGTCGGCGCG GGAGTGCGCG TGCGTGTGA   419

(2) INFORMATION FOR SEQ ID NO:31:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 94 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly
1   5   10   15

Tyr Ala Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys
            20   25   30

Glu Ala Ala Ala Arg Val Tyr Asp Leu Gly Leu Arg Arg Leu Arg Gln
        35   40   45

Arg Arg Arg Leu Arg Arg Glu Arg Val Arg Ala Gln Pro Cys Cys Arg
    50   55   60

Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Ala His Ser Arg
65   70   75   80

Tyr His Thr Val His Glu Leu Ser Ala Arg Glu Cys Ala Cys
                85   90

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:

            (A) LENGTH: 94 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

    Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly
    1   5   10   15

    Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys
                20   25   30

    Glu Ala Ala Ile Arg Ile Tyr Asp Leu Gly Leu Arg Arg Leu Arg Gln
            35   40   45

    Arg Arg Arg Val Arg Arg Glu Arg Ala Arg Ala His Pro Cys Cys Arg
        50   55   60

    Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Val His Ser Arg
    65   70   75   80

    Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu Cys Ala Cys
                85   90

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 8 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 2
            (D) OTHER INFORMATION: /note= "SERINE OR THREONINE"

    (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 3
            (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

    Val Xaa Xaa Leu Gly Leu Gly Tyr
    1   5

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 15 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:

136

```
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 2
                    (D) OTHER INFORMATION: /note= "THREONINE OR GLUTAMIC ACID"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 3
                    (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 4
                    (D) OTHER INFORMATION: /note= "LEUCINE OR ISOLEUCINE"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 9
                    (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 11
                    (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 13
                    (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
    ACID"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 14
                    (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"


        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

        Glu Xaa Xaa Xaa Phe Arg Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Ala
        1               5                   10                  15

(2) INFORMATION FOR SEQ ID NO:35:

        (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 15 amino acids
                    (B) TYPE: amino acid
                    (C) STRANDEDNESS:
                    (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide


        (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 5
                    (D) OTHER INFORMATION: /note= "THREONINE OR VALINE OR
    ISOLEUCINE"

        (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 7
                    (D) OTHER INFORMATION: /note= "TYROSINE OR PHENYLALANINE"

        (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 8
                    (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
    ACID"
```

137

```
    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 10
         (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 11
         (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

    Cys Cys Arg Pro Xaa Ala Xaa Xaa Asp Xaa Xaa Ser Phe Leu Asp
    1               5                   10                  15

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 11 amino acids
         (B) TYPE: amino acid
         (C) STRANDEDNESS:
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 5
         (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 7
         (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 9
         (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 10
         (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

    Phe Arg Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Ala
    1               5                   10

(2) INFORMATION FOR SEQ ID NO:37:

    (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 11 amino acids
         (B) TYPE: amino acid
         (C) STRANDEDNESS:
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (ix) FEATURE:
         (A) NAME/KEY: Modified-site
         (B) LOCATION: 5
```

        (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 7
        (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 9
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 10
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

    Phe Arg Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Ala
    1   5   10

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "ISOLEUCINE OR THREONINE OR
VALINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 7
        (D) OTHER INFORMATION: /note= "TYROSINE OR PHENYLALANINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 10
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

    Cys Cys Arg Pro Xaa Ala Xaa Xaa Asp Xaa
    1   5   10

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid

139

```
          (C) STRANDEDNESS:
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 2
          (D) OTHER INFORMATION: /note= "TYROSINE OR PHENYLALANINE"

    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 3
          (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 5
          (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 6
          (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

    Ala Xaa Xaa Asp Xaa Xaa Ser Phe Leu Asp
    1               5                   10

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 8 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS:
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 2
          (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR THREONINE"

    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 3
          (D) OTHER INFORMATION: /note= "LEUCINE OR VALINE"

    (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 4
          (D) OTHER INFORMATION: /note= "ISOLEUCINE OR LEUCINE"


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

    Glu Xaa Xaa Xaa Phe Arg Tyr Cys
    1               5

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
```

```
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide


   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR THREONINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "LEUCINE OR VALINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 4
        (D) OTHER INFORMATION: /note= "ISOLEUCINE OR LEUCINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 9
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 11
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 13
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"


   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

   Glu Xaa Xaa Xaa Phe Arg Tyr Cys Xaa Gly Xaa Cys Xaa
   1               5                   10

(2) INFORMATION FOR SEQ ID NO:42:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 23 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA



    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GTNWSNGANY TNGGNYTNGG NTA                                         23

(2) INFORMATION FOR SEQ ID NO:43:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 32 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear
```

141

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TTYMGNTAYT GYDSNGGNDS NTGYGANKCN GC   32

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GCNGMNTCRC ANSHNCCNSH RCARTANCKR AA   32

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TCRTCNTCRW ANGCNRYNGG NCKRCARCA   29

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TCNARRAANS WNAVNTCRTC NTCRWANGC   29

(2) INFORMATION FOR SEQ ID NO:47:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GARRMNBTNH TNTTYMGNTA YTG   23

(2) INFORMATION FOR SEQ ID NO:48:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 38 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GARRMNBTNH TNTTYMGNTA YTGYDSNGGN DSNTGHGA   38

(2) INFORMATION FOR SEQ ID NO:49:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 16 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS:
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Ser Gly Ala Arg Pro Xaa Gly Leu Arg Glu Leu Glu Val Ser Val Ser
1   5   10   15

(2) INFORMATION FOR SEQ ID NO:50:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 17 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CCNACNGCNT AYGARGA   17

(2) INFORMATION FOR SEQ ID NO:51:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 20 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS:
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Ala Arg Ala His Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val
1   5   10   15

143

Ser Phe Leu Asp
                20

(2) INFORMATION FOR SEQ ID NO:52:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 20 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ARYTCYTGNA RNGTRTGRTA   20

(2) INFORMATION FOR SEQ ID NO:53:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 28 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GACGAGGTGT CCTTCCTGGA CGTACACA   28

(2) INFORMATION FOR SEQ ID NO:54:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

TAGCGGCTGT GTACGTCCAG GAAGGACACC TCGT   34

(2) INFORMATION FOR SEQ ID NO:55:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CAGCGACGAC GCGTGCGCAA AGAGCG   26

(2) INFORMATION FOR SEQ ID NO:56:

144

```
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 47 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA



       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

TAYGARGACG AGGTGTCCTT CCTGGACGTA CACAGCCGCT AYCAYAC                47

(2) INFORMATION FOR SEQ ID NO:57:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA



       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GCGGCCATCC GCATCTACGA CCTGGG                                      26

(2) INFORMATION FOR SEQ ID NO:58:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA



       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CRTAGGCCGT CGGGCGRCAR CACGGGT                                     27

(2) INFORMATION FOR SEQ ID NO:59:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: cDNA



       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GCGCCGAAGG CCCAGGTCGT AGATGCG                                     27

(2) INFORMATION FOR SEQ ID NO:60:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 29 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear
```

145

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CGCTACTGCG CAGGCGCGTG CGARGCGGC   29

(2) INFORMATION FOR SEQ ID NO:61:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CGCCGACAGC TCTTGCAGCG TRTGGTA   27

(2) INFORMATION FOR SEQ ID NO:62:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GAGCTGGGCC TGGGCTACGC GTCCGACGAG   30

(2) INFORMATION FOR SEQ ID NO:63:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 39 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GCGACGCGTA CCATGAGGCG CTGGAAGGCA GCGGCCCTG   39

(2) INFORMATION FOR SEQ ID NO:64:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

146

GACGGATCCG CATCACACGC ACGCGCACTC   30

(2) INFORMATION FOR SEQ ID NO:65:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 29 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GACCATATGC CGGGGGCTCG GCCTTGTGG   29

(2) INFORMATION FOR SEQ ID NO:66:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GACGGATCCG CATCACACGC ACGCGCACTC   30

(2) INFORMATION FOR SEQ ID NO:67:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CAGCGACGAC GCGTGCGCAA AGAGCG   26

(2) INFORMATION FOR SEQ ID NO:68:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

TAGCGGCTGT GTACGTCCAG GAAGGACACC TCGT   34

(2) INFORMATION FOR SEQ ID NO:69:

        (i) SEQUENCE CHARACTERISTICS:

147

```
                    (A) LENGTH: 21 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

AAAAATCGGG GGTGYGTCTT A.                                              21

(2) INFORMATION FOR SEQ ID NO:70:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 21 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATGCCTGGC CTACYTTGTC A                                               21

(2) INFORMATION FOR SEQ ID NO:71:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 24 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CTGGCGTCCC AMCAAGGGTC TTCG                                            24

(2) INFORMATION FOR SEQ ID NO:72:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 23 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GCCAGTGGTG CCGTCGAGGC GGG                                             23

(2) INFORMATION FOR SEQ ID NO:73:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 24 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear
```

148

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GGCCCAGGAT GAGGCGCTGG AAGG   24

(2) INFORMATION FOR SEQ ID NO:74:

  (i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 27 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CCACTCCACT GCCTGAWATT CWACCCC   27

(2) INFORMATION FOR SEQ ID NO:75:

  (i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 24 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: cDNA

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CCATGTGATT ATCGACCATT CGGC   24

(2) INFORMATION FOR SEQ ID NO:76:

  (i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 134 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: protein

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

  Ser Pro Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg
  1   5   10   15

  Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg
     20   25   30

  Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu
    35   40   45

  Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile
    50   55   60

  Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr Asp
  65   70   75   80

```
      Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp Lys
                    85                  90                      95

      Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser
                    100                 105             110

      Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala
                    115                 120             125

      Lys Arg Cys Gly Cys Ile
                    130
```

(2) INFORMATION FOR SEQ ID NO:77:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 134 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```
      Ser Pro Asp Lys Gln Ala Ala Ala Leu Pro Arg Arg Glu Arg Asn Arg
      1               5                   10                  15

      Gln Ala Ala Ala Ala Ser Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg
                    20                  25                  30

      Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu
                    35                  40                  45

      Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile
            50                  55                  60

      Phe Arg Tyr Cys Ser Gly Ser Cys Glu Ser Ala Glu Thr Met Tyr Asp
      65                  70                  75                  80

      Lys Ile Leu Lys Asn Leu Ser Arg Ser Arg Arg Leu Thr Ser Asp Lys
                    85                  90                  95

      Val Gly Gln Ala Cys Cys Arg Pro Val Ala Phe Asp Asp Asp Leu Ser
                    100                 105             110

      Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala
                    115                 120             125

      Lys Arg Cys Gly Cys Ile
                    130
```

(2) INFORMATION FOR SEQ ID NO:78:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 134 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```
      Ser Pro Asp Lys Gln Ala Ala Ala Leu Pro Arg Arg Glu Arg Asn Arg
      1               5                   10                  15
```

150

```
Gln Ala Ala Ala Ala Ser Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg
        20              25              30

Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu
        35              40              45

Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile
    50              55              60

Phe Arg Tyr Cys Ser Gly Ser Cys Glu Ala Ala Glu Thr Met Tyr Asp
65              70              75              80

Lys Ile Leu Lys Asn Leu Ser Arg Ser Arg Arg Leu Thr Ser Asp Lys
            85              90              95

Val Gly Gln Ala Cys Cys Arg Pro Val Ala Phe Asp Asp Asp Leu Ser
            100             105             110

Phe Leu Asp Asp Ser Leu Val Tyr His Ile Leu Arg Lys His Ser Ala
        115             120             125

Lys Arg Cys Gly Cys Ile
        130
```

(2) INFORMATION FOR SEQ ID NO:79:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 89 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```
Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly
1               5               10              15

Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr Cys Ala Gly Ser Cys
            20              25              30

Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg
        35              40              45

Gly Arg Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr
    50              55              60

Ala Asp Val Thr Phe Leu Asp Asp Gln His His Trp Gln Gln Leu Pro
65              70              75              80

Gln Leu Ser Ala Ala Ala Cys Gly Cys
            85
```

(2) INFORMATION FOR SEQ ID NO:80:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 96 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

151

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Ala Leu Ala Gly Ser Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala
1   5   10   15

Glu Leu Gly Leu Gly Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr
            20   25   30

Cys Ala Gly Ser Cys Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val
        35   40   45

Leu Ala Arg Leu Arg Gly Arg Gly Arg Ala His Gly Arg Pro Cys Cys
    50   55   60

Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp Asp Gln His His
65   70   75   80

Trp Gln Gln Leu Pro Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly
                85   90   95

(2) INFORMATION FOR SEQ ID NO:81:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 134 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Val Arg Ile Pro Gly Gly Leu Pro Thr Pro Gln Phe Leu Leu Ser Lys
1   5   10   15

Pro Ser Leu Cys Leu Thr Ile Leu Leu Tyr Leu Ala Leu Gly Asn Asn
            20   25   30

His Val Arg Leu Pro Arg Ala Leu Ala Gly Ser Cys Arg Leu Trp Ser
        35   40   45

Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly Tyr Ala Ser Glu Glu
    50   55   60

Lys Val Ile Phe Arg Tyr Cys Ala Gly Ser Cys Pro Gln Glu Ala Arg
65   70   75   80

Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg Gly Arg Gly Arg Ala
                85   90   95

His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe
            100   105   110

Leu Asp Asp Gln His His Trp Gln Gln Leu Pro Gln Leu Ser Ala Ala
            115   120   125

Ala Cys Gly Cys Gly Gly
        130

(2) INFORMATION FOR SEQ ID NO:82:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 89 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:

152

        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

    Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly
    1   5   10   15

    Tyr Ala Ser Glu Glu Lys Ile Ile Phe Arg Tyr Cys Ala Gly Ser Cys
            20   25   30

    Pro Gln Glu Val Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg
            35   40   45

    Gly Gln Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr
        50   55   60

    Ala Asp Val Thr Phe Leu Asp Asp His His His Trp Gln Gln Leu Pro
    65   70   75   80

    Gln Leu Ser Ala Ala Ala Cys Gly Cys
                    85

(2) INFORMATION FOR SEQ ID NO:83:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 91 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

    Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly
    1   5   10   15

    Tyr Ala Ser Glu Glu Lys Ile Ile Phe Arg Tyr Cys Ala Gly Ser Cys
            20   25   30

    Pro Gln Glu Val Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg
            35   40   45

    Gly Gln Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr
        50   55   60

    Ala Asp Val Thr Phe Leu Asp Asp His His His Trp Gln Gln Leu Pro
    65   70   75   80

    Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly
                    85   90

(2) INFORMATION FOR SEQ ID NO:84:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 267 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

153

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TGCCGACTGT GGAGCCTGAC CCTACCAGTG GCTGAGCTGG GCCTGGGCTA TGCCTCGGAG   60

GAGAAGGTCA TCTTCCGATA CTGTGCTGGC AGCTGTCCCC AAGAGGCCCG TACCCAGCAC   120

AGTCTGGTAC TGGCCCGGCT TCGAGGGCGG GGTCGAGCCC ATGGCCGACC CTGCTGCCAG   180

CCCACCAGCT. ATGCTGATGT. GACCTTCCTT. GATGATCAGC ACCATTGGCA GCAGCTGCCT   240

CAGCTCTCAG CTGCAGCTTG TGGCTGT   267

(2) INFORMATION FOR SEQ ID NO:85:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 267 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TGCCGGCTGT GGAGCCTGAC CCTACCAGTG GCTGAGCTTG GCCTGGGCTA TGCCTCAGAG   60

GAGAAGATTA TCTTCCGATA CTGTGCTGGC AGCTGTCCCC AAGAGGTCCG TACCCAGCAC   120

AGTCTGGTGC TGGCCCGTCT TCGAGGGCAG GGTCGAGCTC ATGGCAGACC TTGCTGCCAG   180

CCCACCAGCT ATGCTGATGT GACCTTCCTT GATGACCACC ACCATTGGCA GCAGCTGCCT   240

CAGCTCTCAG CCGCAGCTTG TGGCTGT   267

(2) INFORMATION FOR SEQ ID NO:86:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 273 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

TGCCGGCTGT GGAGCCTGAC CCTACCAGTG GCTGAGCTTG GCCTGGGCTA TGCCTCAGAG   60

GAGAAGATTA TCTTCCGATA CTGTGCTGGC AGCTGTCCCC AAGAGGTCCG TACCCAGCAC   120

AGTCTGGTGC TGGCCCGTCT TCGAGGGCAG GGTCGAGCTC ATGGCAGACC TTGCTGCCAG   180

CCCACCAGCT ATGCTGATGT GACCTTCCTT GATGACCACC ACCATTGGCA GCAGCTGCCT   240

CAGCTCTCAG CCGCAGCTTG TGGCTGTGGT GGC   273

(2) INFORMATION FOR SEQ ID NO:87:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 94 amino acids

154

    (B) TYPE: amino acid
    (C) STRANDEDNESS:
    (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: protein

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```
Cys Val Leu Thr Ala Ile His Leu Asn Val Thr Asp Leu Gly Leu Gly
1               5                   10                  15

Tyr Glu Thr Lys Glu Glu Leu Ile Phe Arg Tyr Cys Ser Gly Ser Cys
            20                  25                  30

Glu Ser Ala Glu Thr Met Tyr Asp Lys Ile Leu Lys Asn Leu Ser Arg
        35                  40                  45

Ser Arg Arg Leu Thr Ser Asp Lys Val Gly Gln Ala Cys Cys Arg Pro
    50                  55                  60

Val Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp Asp Asn Leu Val Tyr
65                  70                  75                  80

His Ile Leu Arg Lys His Ser Ala Lys Arg Cys Gly Cys Ile
            85                  90
```

(2) INFORMATION FOR SEQ ID NO:88:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 95 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: protein

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```
Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly
1               5                   10                  15

Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys
            20                  25                  30

Glu Ala Ala Ile Arg Ile Tyr Asp Leu Gly Leu Arg Arg Leu Arg Gln
        35                  40                  45

Arg Arg Arg Val Arg Arg Glu Arg Ala Arg Ala His Pro Cys Cys Arg
    50                  55                  60

Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Val His Ser Arg
65                  70                  75                  80

Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu Cys Ala Cys Val
            85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO:89:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 91 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

155

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly
1   5   10   15

Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr Cys Ala Gly Ser Cys
            20   25   30

Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg
        35   40   45

Gly Arg Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr
    50   55   60

Ala Asp Val Thr Phe Leu Asp Asp Gln His His Trp Gln Gln Leu Pro
65   70   75   80

Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly
                85   90

(2) INFORMATION FOR SEQ ID NO:90:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

TGCCTCAGAG GAGAAGATTA TC   22

(2) INFORMATION FOR SEQ ID NO:91:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Ala Ser Glu Glu Lys Ile Ile
1   5

(2) INFORMATION FOR SEQ ID NO:92:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

156

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile Phe Arg Tyr Cys
1   5   10   15

(2) INFORMATION FOR SEQ ID NO:93:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Leu Gly Leu Gly Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys
1   5   10   15

(2) INFORMATION FOR SEQ ID NO:94:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Leu Gly Leu Gly Tyr Ala Ser Glu Glu Lys Ile Ile Phe Arg Tyr Cys
   1   5   10   15

(2) INFORMATION FOR SEQ ID NO:95:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

AGTCGGGGTT GGGGTATGCC TCA   23

(2) INFORMATION FOR SEQ ID NO:96:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid

157

```
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: cDNA




      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

TATGCCTCAG AGGAGAAGAT TATCTT                                              26

(2) INFORMATION FOR SEQ ID NO:97:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 336 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: cDNA




      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CCTCAGAGGA GAAGATTATC TTCCGATACT GTGCTGGCAG CTGTCCCCAA GAGGTCCGTA        60

CCCAGCACAG TCTGGTGCTG GCCCGTCTTC GAGGGCAGGG TCGAGCTCAT GGCAGACCTT        120

GCTGCCAGCC CACCAGCTAT GCTGATGTGA CCTTCCTTGA TGACCACCAC CATTGGCAGC        180

AGCTGCCTCA GCTCTCAGCC GCAGCTTGTG GCTGTGGTGG CTGAAGGCGG CCAGCCTGGT        240

CTCTCAGAAT CACAAGCAAG AGGCAGCCTT TGAAAGGCTC AGGTGACGTT ATTAGAAACT        300

TGCATAGGAG AAGATTAAGA AGAGAAAGGG GACCTG                                  336

(2) INFORMATION FOR SEQ ID NO:98:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 17 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS:
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein




      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

      Ala Cys Cys Arg Pro Val Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp
      1               5                  10                  15

      Asp


(2) INFORMATION FOR SEQ ID NO:99:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 17 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS:
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein
```

158

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Lys Asp
1   5   10   15

Val

(2) INFORMATION FOR SEQ ID NO:100:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp Asp
1   5   10   15

(2) INFORMATION FOR SEQ ID NO:101:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

TCATCAAGGA AGGTCACATC AGCATA   26

(2) INFORMATION FOR SEQ ID NO:102:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CCACCACAGC CACAAGCTGC GGSTGAGAGC TG   32

(2) INFORMATION FOR SEQ ID NO:103:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid

159

```
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide



    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

    Ala Leu Ala Gly Ser
    1               5

(2) INFORMATION FOR SEQ ID NO:104:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 43 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein



    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

    Val Arg Ile Pro Gly Gly Leu Pro Thr Pro Gln Phe Leu Leu Ser Lys
    1               5                   10                  15

    Pro Ser Leu Cys Leu Thr Ile Leu Leu Tyr Leu Ala Leu Gly Asn Asn
                20                  25                  30

    His Val Arg Leu Pro Arg Ala Leu Ala Gly Ser
            35                  40

(2) INFORMATION FOR SEQ ID NO:105:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 544 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA



    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GAGGGACCTG GACGCCCCAT CAGGGTAAGA ATTCCTGGGG GCCTCCCGAC TCCCCAATTC      60

CTTCTCTCAA AGCCCTCACT TTGCCTTACA ATCCTACTCT ACCTTGCACT AGGTAACAAC     120

CATGTCCGTC TTCCAAGAGC CTTGGCTGGT TCATGCCGAC TGTGGAGCCT GACCCTACCA     180

GTGGCTGAGC TGGGCCTGGG CTATGCCTCG GAGGAGAAGG TCATCTTCCG ATACTGTGCT     240

GGCAGCTGTC CCCAAGAGGC CCGTACCCAG CACAGTCTGG TACTGGCCCG GCTTCGAGGG     300

CGGGGTCGAG CCCATGGCCG ACCCTGCTGC CAGCCCACCA GCTATGCTGA TGTGACCTTC     360

CTTGATGATC AGCACCATTG GCAGCAGCTG CCTCAGCTCT CAGCTGCAGC TTGTGGCTGT     420

GGTGGCTGAA GGAGGCCAGT CTGGTGTCTC AGAATCACAA GCATGAGACA GGCTGGGCTT     480

TGAAAGGCTC AGGTGACATT ACTAGAAATT TGCATAGGTA AAGATAAGAA GGGAAAGGAC     540
```

160

CAGG   544

(2) INFORMATION FOR SEQ ID NO:106:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 336 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

CCTCAGAGGA GAAGATTATC TTCCGATACT GTGCTGGCAG CTGTCCCCAA GAGGTCCGTA   60

CCCAGCACAG TCTGGTGCTG GCCCGTCTTC GAGGGCAGGG TCGAGCTCAT GGCAGACCTT   120

GCTGCCAGCC CACCAGCTAT GCTGATGTGA CCTTCCTTGA TGACCACCAC CATTGGCAGC   180

AGCTGCCTCA GCTCTCAGCC GCAGCTTGTG GCTGTGGTGG CTGAAGGCGG CCAGCCTGGT   240

CTCTCAGAAT CACAAGCAAG AGGCAGCCTT TGAAAGGCTC AGGTGACGTT ATTAGAAACT   300

TGCATAGGAG AAGATTAAGA AGAGAAAGGG GACCTG   336

(2) INFORMATION FOR SEQ ID NO:107:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 391 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

TGCCGGCTGT GGAGCCTGAC CCTACCAGTG GCTGAGCTTG GCCTGGGCTA TGCCTCAGAG   60

GAGAAGATTA TCTTCCGATA CTGTGCTGGC AGCTGTCCCC AAGAGGTCCG TACCCAGCAC   120

AGTCTGGTGC TGGCCCGTCT TCGAGGGCAG GGTCGAGCTC ATGGCAGACC TTGCTGCCAG   180

CCCACCAGCT ATGCTGATGT GACCTTCCTT GATGACCACC ACCATTGGCA GCAGCTGCCT   240

CAGCTCTCAG CCGCAGCTTG TGGCTGTGGT GGCTGAAGGC GGCCAGCCTG GTCTCTCAGA   300

ATCACAAGCA AGAGGCAGCC TTTGAAAGGC TCAGGTGACG TTATTAGAAA CTTGCATAGG   360

AGAAGATTAA GAAGAGAAAG GGGACCTGAT T   391

(2) INFORMATION FOR SEQ ID NO:108:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 8 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein

161

```
      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 2
            (D) OTHER INFORMATION: /note= "SERINE, THREONINE, OR
ALANINE"

      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 3
            (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"


      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

      Val Xaa Xaa Leu Gly Leu Gly Tyr
      1               5

(2) INFORMATION FOR SEQ ID NO:109:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 8 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein


      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 5
            (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 7
            (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"


      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

      Phe Arg Tyr Cys Xaa Gly Xaa Cys
      1               5

(2) INFORMATION FOR SEQ ID NO:110:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 8 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein


      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 2
            (D) OTHER INFORMATION: /note= "ASPARTIC ACID, GLUTAMIC
ACID OR NO AMINO ACID"

      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 3
            (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"

      (ix) FEATURE:
            (A) NAME/KEY: Modified-site
```

162

```
              (B) LOCATION: 4
              (D) OTHER INFORMATION: /note= "SERINE OR THREONINE"

      (ix) FEATURE:
              (A) NAME/KEY: Modified-site
              (B) LOCATION: 8
              (D) OTHER INFORMATION: /note= "valine or aspartic acid"


      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

      Asp Xaa Xaa Xaa Phe Leu Asp Xaa
      1               5

(2) INFORMATION FOR SEQ ID NO:111:

      (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 142 amino acids
              (B) TYPE: amino acid
              (C) STRANDEDNESS:
              (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: protein



      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

      Glu Gly Pro Gly Arg Pro Ile Arg Val Arg Ile Pro Gly Gly Leu Pro
      1               5               10              15

      Thr Pro Gln Phe Leu Leu Ser Lys Pro Ser Leu Cys Leu Thr Ile Leu
                  20              25              30

      Leu Tyr Leu Ala Leu Gly Asn Asn His Val Arg Leu Pro Arg Ala Leu
              35              40              45

      Ala Gly Ser Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu
          50              55              60

      Gly Leu Gly Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr Cys Ala
      65              70              75              80

      Gly Ser Cys Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val Leu Ala
                  85              90              95

      Arg Leu Arg Gly Arg Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro
                  100             105             110

      Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp Asp Gln His His Trp Gln
              115             120             125

      Gln Leu Pro Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly
          130             135             140

(2) INFORMATION FOR SEQ ID NO:112:

      (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 5 amino acids
              (B) TYPE: amino acid
              (C) STRANDEDNESS:
              (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: peptide
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Ala Leu Pro Gly Leu
1   5

(2) INFORMATION FOR SEQ ID NO:113:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 12 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 2
            (D) OTHER INFORMATION: /note= "THREONINE, GLUTAMIC ACID OR
LYSINE"

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 3
            (D) OTHER INFORMATION: /note= "VALINE, LEUCINE OR
ISOLEUCINE"

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 4
            (D) OTHER INFORMATION: /note= "LEUCINE OR ISOLEUCINE"

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 9
            (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 11
            (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Glu Xaa Xaa Xaa Phe Arg Tyr Cys Xaa Gly Xaa Cys
1   5   10

(2) INFORMATION FOR SEQ ID NO:114:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 16 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 3
            (D) OTHER INFORMATION: /note= "ARGININE OR GLUTAMINE"

        (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 5
            (D) OTHER INFORMATION: /note= "THREONINE, VALINE OR

ISOLEUCINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "ALANINE OR SERINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 7
        (D) OTHER INFORMATION: /note= "TYROSINE OR PHENYLALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID, ASPARTIC
ACID OR ALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 10
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID, ASPARTIC
ACID OR NO AMINO ACID"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 11
        (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 12
        (D) OTHER INFORMATION: /note= "SERINE OR THREONINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 16
        (D) OTHER INFORMATION: /note= "ASPARTIC ACID OR VALINE"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

   Cys Cys Xaa Pro Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Phe Leu Asp Xaa
   1   5   10   15

(2) INFORMATION FOR SEQ ID NO:115:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

GTNDGNGANY TGGGNYTGGG NTA   23

(2) INFORMATION FOR SEQ ID NO:116:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single

165

```
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GANBTNWCNT TYYTNGANG                                              19

(2) INFORMATION FOR SEQ ID NO:117:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 20 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

GANBTNWCNT TYYTNGANGW                                            20

(2) INFORMATION FOR SEQ ID NO:118:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 23 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

TTYMGNTAYT GYDSNGGNDS NTG                                        23

(2) INFORMATION FOR SEQ ID NO:119:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 20 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA



          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

GTNDGNGANY TGGGNYTNGG                                            20

(2) INFORMATION FOR SEQ ID NO:120:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 23 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: single
                    (D) TOPOLOGY: linear
```

166

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

GTNDGNGANY TGGGNYTGGG NTT   23

(2) INFORMATION FOR SEQ ID NO:121:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

WCNTCNARRA ANGWNAVNTC   20

(2) INFORMATION FOR SEQ ID NO:122:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

WCNTCNARRA ANGWNAVNT   19

(2) INFORMATION FOR SEQ ID NO:123:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

CANSHNCCNS HRCARTANCK RAA   23

(2) INFORMATION FOR SEQ ID NO:124:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

CANSHNCCNS HRCARTANCK RAANA   25

(2) INFORMATION FOR SEQ ID NO:125:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "THREONINE, SERINE OR
ALANINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

    Val Xaa Xaa Leu Gly Leu Gly Tyr
    1   5

(2) INFORMATION FOR SEQ ID NO:126:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1
        (D) OTHER INFORMATION: /note= "ASPARTIC ACID OR GLUTAMIC
ACID"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "VALINE OR LEUCINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "THREONINE OR SERINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "ASPARTIC ACID OR GLUTAMIC
ACID"

168

```
(ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 7
     (D) OTHER INFORMATION: /note= "ASPARTIC ACID OR VALINE"


(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

 Xaa Xaa Xaa Phe Leu Xaa Xaa
 1               5
```

(2) INFORMATION FOR SEQ ID NO:127:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein


   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 7
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"


   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

    Phe Arg Tyr Cys Xaa Gly Xaa Cys
    1               5
```

(2) INFORMATION FOR SEQ ID NO:128:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein


   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "THREONINE, SERINE OR
ALANINE"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "ASPARTIC ACID OR GLUTAMIC
ACID"


   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

    Val Xaa Xaa Leu Gly Leu Gly
    1               5
```

(2) INFORMATION FOR SEQ ID NO:129:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "THREONINE, SERINE OR
ALANINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "GLUTAMIC ACID OR ASPARTIC
ACID"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

    Val Xaa Xaa Leu Gly Leu Gly Phe
    1   5

(2) INFORMATION FOR SEQ ID NO:130:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1
        (D) OTHER INFORMATION: /note= "ISOLEUCINE OR LEUCINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "SERINE OR ALANINE"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

    Xaa Phe Arg Tyr Cys Xaa Gly Xaa Cys
    1   5

(2) INFORMATION FOR SEQ ID NO:131:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 559 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

170

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

ATGGCTGCAG	GAAGACTTCG	GATCCTGTGT	CTGCTGCTCC	TGTCCTTGCA	CCCGAGCCTC	60
GGCTGGGTCC	TTGATCTTCA	AGAGGCTTCT	GTGGCAGATA	AGCTCTCATT	TGGGAAGATG	120
GCAGAGACTA	GAGGGACCTG	GACGCCCCAT	CAGGGTAAGA	ATTCCTGGGG	GCCTCCCGAC	180
TCCCCAATTC	CTTCTCTCAA	AGCCCTCACT	TTGCCTTACA	ATCCTACTCT	ACCTTGCACT	240
AGGTAACAAC	CATGTCCGTC	TTCCAAGAGC	CTTGGCTGGT	TCATGCCGAC	TGTGGAGCCT	300
GACCCTACCA	GTGGCTGAGC	TGGGCCTGGG	CTATGCCTCG	GAGGAGAAGG	TCATCTTCCG	360
ATACTGTGCT	GGCAGCTGTC	CCCAAGAGGC	CCGTACCCAG	CACAGTCTGG	TACTGGCCCG	420
GCTTCGAGGG	CGGGGTCGAG	CCCATGGCCG	ACCCTGCTGC	CAGCCCACCA	GCTATGCTGA	480
TGTGACCTTC	CTTGATGATC	AGCACCATTG	GCAGCAGCTG	CCTCAGCTCT	CAGCTGCAGC	540
TTGTGGCTGT	GGTGGCTGA					559

(2) INFORMATION FOR SEQ ID NO:132:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 81 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Met	Ala	Ala	Gly	Arg	Leu	Arg	Ile	Leu	Cys	Leu	Leu	Leu	Leu	Ser	Leu
1			5					10					15	
His	Pro	Ser	Leu	Gly	Trp	Val	Leu	Asp	Leu	Gln	Glu	Ala	Ser	Val	Ala
			20				25					30		
Asp	Lys	Leu	Ser	Phe	Gly	Lys	Met	Ala	Glu	Thr	Arg	Gly	Thr	Trp	Thr
		35				40					45			
Pro	His	Gln	Gly	Lys	Asn	Ser	Trp	Gly	Pro	Pro	Asp	Ser	Pro	Ile	Pro
	50				55				60				
Ser	Leu	Lys	Ala	Leu	Thr	Leu	Pro	Tyr	Asn	Pro	Thr	Leu	Pro	Cys	Thr
65				70				75					80
Arg														

(2) INFORMATION FOR SEQ ID NO:133:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 185 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

171

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

```
Trp Leu Gln Glu Asp Phe Gly Ser Cys Val Cys Cys Ser Cys Pro Cys
1               5                   10              15

Thr Arg Ala Ser Ala Gly Ser Leu Ile Phe Lys Arg Leu Leu Trp Gln
            20              25              30

Ile Ser Ser His Leu Gly Arg Trp Gln Arg Leu Glu Gly Pro Gly Arg
        35              40              45

Pro Ile Arg Val Arg Ile Pro Gly Gly Leu Pro Thr Pro Gln Phe Leu
    50              55              60

Leu Ser Lys Pro Ser Leu Cys Leu Thr Ile Leu Leu Tyr Leu Ala Leu
65              70              75              80

Gly Asn Asn His Val Arg Leu Pro Arg Ala Leu Ala Gly Ser Cys Arg
            85              90                  95

Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly Tyr Ala
            100             105             110

Ser Glu Glu Lys Val Ile Phe Arg Tyr Cys Ala Gly Ser Cys Pro Gln
        115             120             125

Glu Ala Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg Gly Arg
    130             135             140

Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp
145             150             155             160

Val Thr Phe Leu Asp Asp Gln His His Trp Gln Gln Leu Pro Gln Leu
            165             170             175

Ser Ala Ala Ala Cys Gly Cys Gly Gly
            180             185
```

(2) INFORMATION FOR SEQ ID NO:134:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 559 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

```
ATGGCTGCAG GAAGACTTCG GATCTTGTTT CTGCTGCTCC TGTCCTTGCA CCTGGGCCTT      60

GGCTGGGTCC TTGATCTTCA AGAGGCTCCT GCGGCAGATG AGCTCTCATC TGGGAAAATG     120

GCAGAGACTG GAAGGACCTG GAAGCCCCAT CAGGGTAAGA ATTCTTGGGG GCCTCCTAAC     180

TCTACAGTTC TTCCTCTCAA AGCCCTCACT TTGCCTCACA ATCCTATTCT ACCTTGCACT     240

AGGTAACAAC AATGTCCGCC TTCCAAGAGC CTTACCTGGT TTGTGCCGGC TGTGGAGCCT     300

GACCCTACCA GTGGCTGAGC TTGGCCTGGG CTATGCCTCA GAGGAGAAGA TTATCTTCCG     360

ATACTGTGCT GGCAGCTGTC CCCAAGAGGT CCGTACCCAG CACAGTCTGG TGCTGGCCCG     420

TCTTCGAGGG CAGGGTCGAG CTCATGGCAG ACCTTGCTGC CAGCCCACCA GCTATGCTGA     480

TGTGACCTTC CTTGATGACC ACCACCATTG GCAGCAGCTG CCTCAGCTCT CAGCCGCAGC     540
```

TTGTGGCTGT GGTGGCTGA　　　　　　　　　　　　　　　　　　　　　　　　　　　559

(2) INFORMATION FOR SEQ ID NO:135:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 81 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

```
Met Ala Ala Gly Arg Leu Arg Ile Leu Phe Leu Leu Leu Leu Ser Leu
1               5                   10                  15
His Leu Gly Leu Gly Trp Val Leu Asp Leu Gln Glu Ala Pro Ala Ala
            20                  25                  30
Asp Glu Leu Ser Ser Gly Lys Met Ala Glu Thr Gly Arg Thr Trp Lys
            35                  40                  45
Pro His Gln Gly Lys Asn Ser Trp Gly Pro Pro Asn Ser Thr Val Leu
        50                  55                  60
Pro Leu Lys Ala Leu Thr Leu Pro His Asn Pro Ile Leu Pro Cys Thr
65                  70                  75                  80
Arg
```

(2) INFORMATION FOR SEQ ID NO:136:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 185 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

```
Trp Leu Gln Glu Asp Phe Gly Ser Cys Phe Cys Cys Ser Cys Pro Cys
1               5                   10                  15
Thr Trp Ala Leu Ala Gly Ser Leu Ile Phe Lys Arg Leu Leu Arg Gln
            20                  25                  30
Met Ser Ser His Leu Gly Lys Trp Gln Arg Leu Glu Gly Pro Gly Ser
            35                  40                  45
Pro Ile Arg Val Arg Ile Leu Gly Gly Leu Leu Thr Leu Gln Phe Phe
        50                  55                  60
Leu Ser Lys Pro Ser Leu Cys Leu Thr Ile Leu Phe Tyr Leu Ala Leu
65                  70                  75                  80
Gly Asn Asn Asn Val Arg Leu Pro Arg Ala Leu Pro Gly Leu Cys Arg
                85                  90                  95
Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly Tyr Ala
                100                 105                 110
```

173

```
        Ser Glu Glu Lys Ile Ile Phe Arg Tyr Cys Ala Gly Ser Cys Pro Gln
            115                 120                 125

        Glu Val Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg Gly Gln
            130                 135                 140

        Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp
        145                 150                 155                 160

        Val Thr Phe Leu Asp Asp His His His Trp Gln Gln Leu Pro Gln Leu
                        165             170                 175

        Ser Ala Ala Ala Cys Gly Cys Gly Gly
                    180             185
```

(2) INFORMATION FOR SEQ ID NO:137:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 23 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

AATCCCCAGG ACAGGCAGGG AAT   23

(2) INFORMATION FOR SEQ ID NO:138:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 35 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

CGGTACCCAG ATCTTCAGCC ACCACAGCCA CAAGC   35

(2) INFORMATION FOR SEQ ID NO:139:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 76 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

GGACTATCAT ATGGCCCACC ACCACCACCA CCACCACCAC GACGACGACG ACAAGGCCTT   60

GGCTGGTTCA TGCCGA   76

(2) INFORMATION FOR SEQ ID NO:140:

174

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 35 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

CGGTACCCAG ATCTTCAGCC ACCACAGCCA CAAGC   35

(2) INFORMATION FOR SEQ ID NO:141:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 96 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

Ala Leu Ala Gly Ser Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala
1   5   10   15

Glu Leu Gly Leu Gly Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr
        20   25   30

Cys Ala Gly Ser Cys Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val
        35   40   45

Leu Ala Arg Leu Arg Gly Arg Gly Arg Ala His Gly Arg Pro Cys Cys
    50   55   60

Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Val His Ser
65   70   75   80

Arg Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu Cys Ala Cys Val
            85   90   95

(2) INFORMATION FOR SEQ ID NO:142:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

TAATACGACT CACTATAGGG GAA   23

(2) INFORMATION FOR SEQ ID NO:143:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 49 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single

175

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

TCGTCTTCGT AAGCAGTCGG ACGGCAGCAG GGTCGGCCAT GGGCTCGAC   49

(2) INFORMATION FOR SEQ ID NO:144:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

TGCTGCCGTC CGACTGCTTA CGAAGACGA   29

(2) INFORMATION FOR SEQ ID NO:145:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

GTTATGCTAG TTATTGCTCA GCGGT   25

(2) INFORMATION FOR SEQ ID NO:146:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 100 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

Pro Gly Ala Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser
1   5   10   15

Glu Leu Gly Leu Gly Tyr Thr Ser Asp Glu Thr Val Leu Phe Arg Tyr
        20   25   30

Cys Ala Gly Ala Cys Glu Ala Ala Ile Arg Ile Tyr Asp Leu Gly Leu
        35   40   45

176

```
Arg Arg Leu Arg Gln Arg Arg Arg Val Arg Arg Glu Arg Ala Arg Ala
    50              55              60
```

```
His Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp
65              70              75              80
```

```
Asp Gln His His Trp Gln Gln Leu Pro Gln Leu Ser Ala Ala Ala Cys
                85              90              95
```

```
Gly Cys Gly Gly
        100
```

(2) INFORMATION FOR SEQ ID NO:147:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 50 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CACATCAGCA TAGCTGGTGG GCTGGCAGCA CGGGTGAGCA CGAGCACGTT   50

(2) INFORMATION FOR SEQ ID NO:148:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 25 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

TGCTGCCAGC CCACCAGCTA TGCTG   25

(2) INFORMATION FOR SEQ ID NO:149:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 23 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CCTCGGAGGA GAAGGTCATC TTC   23

(2) INFORMATION FOR SEQ ID NO:150:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 98 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single

177

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

Cys Cys Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp
1   5   10   15

Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly
            20   25   30

Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu
        35   40   45

Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys
        50   55   60

Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg
65   70   75   80

Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys
                85   90   95
Cys Ser

(2) INFORMATION FOR SEQ ID NO:151:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 98 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Cys Cys Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp
1   5   10   15

Lys Trp Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly
            20   25   30

Ala Cys Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu
        35   40   45

Ser Leu Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys
        50   55   60

Val Ser Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys
65   70   75   80

Thr Pro Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys
                85   90   95

Cys Ser

(2) INFORMATION FOR SEQ ID NO:152:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 98 amino acids

178

```
        (B)  TYPE: amino acid
        (C)  STRANDEDNESS: single
        (D)  TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide



    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

    Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp
    1               5                   10                  15

    Lys Trp Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly
                20                  25                  30

    Pro Cys Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu
                35              40                  45

    Gly Leu Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys
            50                  55              60

    Val Pro Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg
    65                  70                  75                  80

    Thr Pro Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys
                    85                  90                  95

    Cys Ser
```

(2) INFORMATION FOR SEQ ID NO:153:

```
    (i)  SEQUENCE CHARACTERISTICS:
        (A)  LENGTH: 106 amino acids
        (B)  TYPE: amino acid
        (C)  STRANDEDNESS: single
        (D)  TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

    Cys Cys Lys Lys Gln Phe Phe Val Ser Phe Lys Asp Ile Gly Trp Asn
    1               5                   10                  15

    Asp Trp Ile Ile Ala Pro Ser Gly Tyr His Ala Asn Tyr Cys Glu Gly
                20                  25                  30

    Glu Cys Pro Ser His Ile Ala Gly Thr Ser Gly Ser Ser Leu Ser Phe
                35              40                  45

    His Ser Thr Val Ile Asn His Tyr Arg Met Arg Gly His Ser Pro Phe
            50                  55              60

    Ala Asn Leu Lys Ser Cys Cys Val Pro Thr Lys Leu Arg Pro Met Ser
    65                  70                  75                  80

    Met Leu Tyr Tyr Asp Asp Gly Gln Asn Ile Ile Lys Lys Asp Ile Gln
                    85                  90                  95

    Asn Met Ile Val Glu Glu Cys Gly Cys Ser
                100                 105
```

(2) INFORMATION FOR SEQ ID NO:154:

```
    (i)  SEQUENCE CHARACTERISTICS:
```

        (A) LENGTH: 105 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

```
Cys Cys Arg Gln Gln Phe Phe Ile Asp Phe Arg Leu Ile Gly Trp Asn
1               5                   10                  15

Asp Trp Ile Ile Ala Pro Thr Gly Tyr Tyr Gly Asn Tyr Cys Glu Gly
            20              25              30

Ser Cys Pro Ala Tyr Leu Ala Gly Val Pro Gly Ser Ala Ser Ser Phe
        35              40              45

His Thr Ala Val Val Asn Gln Tyr Arg Met Arg Gly Leu Asn Pro Gly
    50              55              60

Thr Val Asn Ser Cys Cys Ile Pro Thr Lys Leu Ser Thr Met Ser Met
65              70              75                  80

Leu Tyr Phe Asp Asp Glu Tyr Asn Ile Val Lys Arg Asp Val Pro Asn
            85              90                  95

Met Ile Val Glu Glu Cys Gly Cys Ala
        100             105
```

(2) INFORMATION FOR SEQ ID NO:155:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 101 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

```
Cys Arg Arg Val Lys Phe Gln Val Asp Phe Asn Leu Ile Gly Trp Gly
1               5                   10                  15

Ser Trp Ile Ile Tyr Pro Lys Gln Tyr Asn Ala Tyr Arg Cys Glu Gly
            20              25              30

Glu Cys Pro Asn Pro Val Gly Glu Glu Phe His Pro Thr Asn His Ala
        35              40              45

Tyr Ile Gln Ser Leu Leu Lys Arg Tyr Gln Pro His Arg Val Pro Ser
    50              55              60

Thr Cys Cys Ala Pro Val Lys Thr Lys Pro Leu Ser Met Leu Tyr Val
65              70              75                  80

Asp Asn Gly Arg Val Leu Leu Glu His His Lys Asp Met Ile Val Glu
            85              90                  95

Glu Cys Gly Cys Leu
        100
```

180.

(2) INFORMATION FOR SEQ ID NO:156:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 101 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

```
Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
1               5                   10                  15

Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly
            20                  25                  30

Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
        35                  40                  45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala
    50                  55                  60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
65                  70                  75                  80

Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu
                85                  90                  95

Gly Cys Gly Cys Arg
                100
```

(2) INFORMATION FOR SEQ ID NO:157:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 101 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

```
Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
1               5                   10                  15

Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly
            20                  25                  30

Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
        35                  40                  45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys Ala
    50                  55                  60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
65                  70                  75                  80

Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu
                85                  90                  95
```

181

```
Gly Cys Gly Cys Arg
            100
```

(2) INFORMATION FOR SEQ ID NO:158:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 102 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

```
Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asp
1               5                   10                  15

Asp Trp Ile Val Ala Pro Leu Gly Tyr Asp Ala Tyr Tyr Cys His Gly
            20              25              30

Lys Cys Pro Phe Pro Leu Ala Asp His Phe Asn Ser Thr Asn His Ala
        35              40              45

Val Val Gln Thr Leu Val Asn Asn Met Asn Pro Gly Lys Val Pro Lys
    50              55              60

Ala Cys Cys Val Pro Thr Gln Leu Asp Ser Val Ala Met Leu Tyr Leu
65              70              75              80

Asn Asp Gln Ser Thr Val Val Leu Lys Asn Tyr Gln Glu Met Thr Val
            85              90              95

Val Gly Cys Gly Cys Arg
            100
```

(2) INFORMATION FOR SEQ ID NO:159:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 102 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

```
Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
1               5                   10                  15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly
            20              25              30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
        35              40              45

Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys
    50              55              60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65              70              75              80
```

182

```
Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
            85                    90                    95

Arg Ser Cys Gly Cys His
            100
```

(2) INFORMATION FOR SEQ ID NO:160:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

```
Cys Arg Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln
1                5                    10                    15

Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly
            20                    25                    30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
            35                    40                    45

Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro Lys
            50                    55                    60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65                    70                    75                    80

Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
            85                    90                    95

Arg Ala Cys Gly Cys His
            100
```

(2) INFORMATION FOR SEQ ID NO:161:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

```
Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
1                5                    10                    15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly
            20                    25                    30

Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr Asn His Ala
            35                    40                    45

Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr Val Pro Lys
            50                    55                    60
```

183

```
Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val Leu Tyr Phe
65              70                  75                      80

Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
            85                  90                  95

Arg Ala Cys Gly Cys His
            100
```

(2) INFORMATION FOR SEQ ID NO:162:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

```
Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu
1               5                   10                  15

Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly
            20                  25                  30

Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala
        35                  40                  45

Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys
    50                  55                  60

Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr
65              70                  75                      80

Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val
            85                  90                  95

Lys Ala Cys Gly Cys His
            100
```

(2) INFORMATION FOR SEQ ID NO:163:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

```
Cys Gln Met Gln Thr Leu Tyr Ile Asp Phe Lys Asp Leu Gly Trp His
1               5                   10                  15
Asp Trp Ile Ile Ala Pro Glu Gly Tyr Gly Ala Phe Tyr Cys Ser Gly
            20                  25                  30

Glu Cys Asn Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
        35                  40                  45

Ile Val Gln Thr Leu Val His Leu Leu Glu Pro Lys Lys Val Pro Lys
    50                  55                  60
```

184

Pro Cys Cys Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr His
65   70   75   80

Leu Asn Asp Glu Asn Val Asn Leu Lys Lys Tyr Arg Asn Met Ile Val
              85   90   95

Lys Ser Cys Gly Cys His
          100

(2) INFORMATION FOR SEQ ID NO:164:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 103 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

    Cys Ala Arg Arg Tyr Leu Lys Val Asp Phe Ala Asp Ile Gly Trp Ser
    1   5   10   15

    Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp Ala Tyr Tyr Cys Ser Gly
            20   25   30

    Ala Cys Gln Phe Pro Met Pro Lys Ser Leu Lys Pro Ser Asn His Ala
            35   40   45

    Thr Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile Pro
        50   55   60

    Glu Pro Cys Cys Val Pro Glu Lys Met Ser Ser Leu Ser Ile Leu Phe
    65   70   75   80

    Phe Asp Glu Asn Lys Asn Val Val Leu Lys Val Tyr Pro Asn Met Thr
                85   90   95

    Val Glu Ser Cys Ala Cys Arg
              100

(2) INFORMATION FOR SEQ ID NO:165:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 102 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

    Cys Lys Lys Arg His Leu Tyr Val Glu Phe Lys Asp Val Gly Trp Gln
    1   5   10   15

    Asn Trp Val Ile Ala Pro Gln Gly Tyr Met Ala Asn Tyr Cys Tyr Gly
            20   25   30

    Glu Cys Pro Tyr Pro Leu Thr Glu Ile Leu Asn Gly Ser Asn His Ala
            35   40   45

185

```
Ile Leu Gln Thr Leu Val His Ser Ile Glu Pro Glu Asp Ile Pro Leu
    50                  55                  60
```

```
Pro Cys Cys Val Pro Thr Lys Met Ser Pro Ile Ser Met Leu Phe Tyr
65                  70                  75                  80
```

```
Asp Asn Asn Asp Asn Val Val Leu Arg His Tyr Glu Asn Met Ala Val
                85                  90                  95
```

```
Asp Glu Cys Gly Cys Arg
            100
```

(2) INFORMATION FOR SEQ ID NO:166:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 106 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

```
Cys Arg Ala Arg Arg Leu Tyr Val Ser Phe Arg Glu Val Gly Trp His
1               5               10                  15
```

```
Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Tyr Cys Gln Gly
            20                  25                  30
```

```
Gln Cys Ala Leu Pro Val Ala Leu Ser Gly Ser Gly Gly Pro Pro Ala
            35                  40                  45
```

```
Leu Asn His Ala Val Leu Arg Ala Leu Met His Ala Ala Ala Pro Gly
    50                  55                  60
```

```
Ala Ala Asp Leu Pro Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser
65                  70                  75                  80
```

```
Val Leu Phe Phe Asp Asn Ser Asp Asn Val Val Leu Arg Gln Tyr Glu
                85                  90                  95
```

```
Asp Met Val Val Asp Glu Cys Gly Cys Arg
            100                 105
```

(2) INFORMATION FOR SEQ ID NO:167:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 101 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

```
Cys His Arg His Gln Leu Phe Ile Asn Phe Gln Asp Leu Gly Trp His
1               5               10                  15
```

```
Lys Trp Val Ile Ala Pro Lys Gly Phe Met Ala Asn Tyr Cys His Gly
            20                  25                  30
```

Glu Cys Pro Phe Ser Met Thr Thr Tyr Leu Asn Ser Ser Asn Tyr Ala
        35   40   45

Phe Met Gln Ala Leu Met His Met Ala Asp Pro Lys Val Pro Lys Ala
    50   55   60

Val Cys Val Pro Thr Lys Leu Ser Pro Ile Ser Met Leu Tyr Gln Asp
65   70   75   80

Ser Asp Lys Asn Val Ile Leu Arg His Tyr Glu Asp Met Val Val Asp
                85   90   95
Glu Cys Gly Cys Gly
            100

(2) INFORMATION FOR SEQ ID NO:168:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 103 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Cys Arg Arg Thr Ser Leu His Val Asn Phe Lys Glu Ile Gly Trp Asp
1   5   10   15

Ser Trp Ile Ile Ala Pro Lys Asp Tyr Glu Ala Phe Glu Cys Lys Gly
            20   25   30

Gly Cys Phe Phe Pro Leu Thr Asp Asn Val Thr Pro Thr Lys His Ala
    35   40   45

Ile Val Gln Thr Leu Val His Leu Gln Asn Pro Lys Lys Ala Ser Lys
    50   55   60

Ala Cys Cys Val Pro Thr Lys Leu Asp Ala Ile Ser Ile Leu Tyr Lys
65   70   75   80

Asp Asp Ala Gly Val Pro Thr Leu Ile Tyr Asn Tyr Glu Gly Met Lys
                85   90   95

Val Ala Glu Cys Gly Cys Arg
            100

(2) INFORMATION FOR SEQ ID NO:169:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 105 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

Cys His Arg Val Ala Leu Asn Ile Ser Phe Gln Glu Leu Gly Trp Glu
1   5   10   15

187

```
Arg Trp Ile Val Tyr Pro Pro Ser Phe Ile Phe His Tyr Cys His Gly
          20              25              30

Gly Cys Gly Leu His Ile Pro Pro Asn Leu Ser Leu Pro Val Pro Gly
          35              40              45

Ala Pro Pro Thr Pro Ala Gln Pro Tyr Ser Leu Leu Pro Gly Ala Gln
    50              55              60

Pro Cys Cys Ala Ala Leu Pro Gly Thr Met Arg Pro Leu His Val Arg
65              70              75              80

Thr Thr Ser Asp Gly Gly Tyr Ser Phe Lys Tyr Glu Thr Val Pro Asn
              85              90              95

Leu Leu Thr Gln His Cys Ala Cys Ile
          100             105
```

(2) INFORMATION FOR SEQ ID NO:170:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 99 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

```
Cys Ala Leu Arg Glu Leu Ser Val Asp Leu Arg Ala Glu Arg Ser Val
1                 5               10              15

Leu Ile Pro Glu Thr Tyr Gln Ala Asn Asn Cys Gln Gly Ala Cys Gly
          20              25              30

Trp Pro Gln Ser Asp Arg Asn Pro Arg Tyr Gly Asn His Val Val Leu
          35              40              45

Leu Leu Lys Met Gln Ala Arg Gly Ala Thr Leu Ala Arg Pro Pro Cys
    50              55              60

Cys Val Pro Thr Ala Tyr Thr Gly Lys Leu Leu Ile Ser Leu Ser Glu
65              70              75              80

Glu Arg Ile Ser Ala His His Val Pro Asn Met Val Ala Thr Glu Cys
              85              90              95

Gly Cys Arg
```

(2) INFORMATION FOR SEQ ID NO:171:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 102 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

```
Cys Glu Leu His Asp Phe Ser Leu Ser Phe Ser Gln Leu Lys Trp Asp
1                 5               10              15
```

188

```
Asn Trp Ile Val Ala Pro His Ser Tyr Asn Pro Ser Tyr Cys Lys Gly
        20              25              30

Asp Cys Pro Ser Ala Val Ser His Arg Tyr Gly Ser Pro Val His Thr
        35              40              45

Met Val Gln Asn Met Ile Tyr Glu Lys Leu Asp Pro Ser Val Pro Ser
50              55              60

Pro Ser Cys Val Pro Gly Lys Tyr Ser Pro Leu Ser Val Leu Thr Ile
65              70              75              80

Glu Pro Asp Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Met Ala
            85              90              95

Thr Ser Cys Thr Cys Arg
            100
```

(2) INFORMATION FOR SEQ ID NO:172:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 94 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

```
Cys Val Leu Thr Ala Ile His Leu Asn Val Thr Asp Leu Gly Leu Gly
1               5               10              15

Tyr Glu Thr Lys Glu Glu Leu Ile Phe Arg Tyr Cys Ser Gly Ser Cys
        20              25              30

Asp Ala Ala Glu Thr Thr Tyr Asp Lys Ile Leu Lys Asn Leu Ser Arg
        35              40              45

Asn Arg Arg Leu Val Ser Asp Lys Val Gly Gln Ala Cys Cys Arg Pro
        50              55              60

Ile Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp Asp Asn Leu Val Tyr
65              70              75              80

His Ile Leu Arg Lys His Ser Ala Lys Arg Cys Gly Cys Ile
            85              90
```

(2) INFORMATION FOR SEQ ID NO:173:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 95 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

```
Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly
1               5               10              15
```

189

```
Tyr Ala Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys
            20                  25                  30

Glu Ala Ala Ala Arg Val Tyr Asp Leu Gly Leu Arg Arg Leu Arg Gln
        35                  40                  45

Arg Arg Arg Leu Arg Arg Glu Arg Val Arg Ala Gln Pro Cys Cys Arg
    50                  55                  60

Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Ala His Ser Arg
65                  70                  75                  80

Tyr His Thr Val His Glu Leu Ser Ala Arg Glu Cys Ala Cys Val
                85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO:174:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 291 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

```
GCCTTGGCTG GTTCATGCCG ACTGTGGAGC CTGACCCTAC CAGTGGCTGA GCTGGGCCTG      60

GGCTATGCCT CGGAGGAGAA GGTCATCTTC CGATACTGTG CTGGCAGCTG TCCCCAAGAG     120

GCCCGTACCC AGCACAGTCT GGTACTGGCC CGGCTTCGAG GGCGGGGTCG AGCCCATGGC     180

CGACCCTGCT GCCAGCCCAC CAGCTATGCT GATGTGACCT TCCTTGATGA TCAGCACCAT     240

TGGCAGCAGC TGCCTCAGCT CTCAGCTGCA GCTTGTGGCT GTGGTGGCTG A              291
```

(2) INFORMATION FOR SEQ ID NO:175:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 405 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

```
GTAAGAATTC CTGGGGGCCT CCCGACTCCC CAATTCCTTC TCTCAAAGCC CTCACTTTGC      60

CTTACAATCC TACTCTACCT TGCACTAGGT AACAACCATG TCCGTCTTCC AAGAGCCTTG     120

GCTGGTTCAT GCCGACTGTG GAGCCTGACC CTACCAGTGG CTGAGCTGGG CCTGGGCTAT     180

GCCTCGGAGG AGAAGGTCAT CTTCCGATAC TGTGCTGGCA GCTGTCCCCA AGAGGCCCGT     240

ACCCAGCACA GTCTGGTACT GGCCCGGCTT CGAGGGCGGG GTCGAGCCCA TGGCCGACCC     300

TGCTGCCAGC CCACCAGCTA TGCTGATGTG ACCTTCCTTG ATGATCAGCA CCATTGGCAG     360

CAGCTGCCTC AGCTCTCAGC TGCAGCTTGT GGCTGTGGTG GCTGA                     405
```

(2) INFORMATION FOR SEQ ID NO:176:

190

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 291 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

GCCTTACCTG GTTTGTGCCG GCTGTGGAGC CTGACCCTAC CAGTGGCTGA GCTTGGCCTG   60

GGCTATGCCT CAGAGGAGAA GATTATCTTC CGATACTGTG CTGGCAGCTG TCCCCAAGAG   120

GTCCGTACCC AGCACAGTCT GGTGCTGGCC CGTCTTCGAG GGCAGGGTCG AGCTCATGGC   180

AGACCTTGCT GCCAGCCCAC CAGCTATGCT GATGTGACCT TCCTTGATGA CCACCACCAT   240

TGGCAGCAGC TGCCTCAGCT CTCAGCCGCA GCTTGTGGCT GTGGTGGCTG A   291

75875

191

What is Claimed is:

1. An isolated and purified growth factor comprising persephin.

2. The isolated and purified growth factor of claim 1 comprising a polypeptide sequence which is at least about 85% sequence identity with SEQ ID NO:79 or SEQ ID NO:82.

3. The isolated and purified growth factor of claim 2 comprising a polypeptide sequence selected from the group consisting of SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:83 and conservatively substituted variants thereof.

4. The isolated and purified growth factor of Claim 1 in a pharmaceutically acceptable carrier.

5. The isolated and purified growth factor of claim 1 comprising a polypeptide containing a sequence of amino acids having at least about 65% sequence identity with SEQ ID NO:79 or SEQ ID NO:82 wherein the growth factor is from a non-mammalian species.

6. A method for obtaining a neurturin-persephin-GDNF family member growth factor comprising:

(a) isolating from a human genomic or cDNA library, a clone that hybridizes with a polynucleotide
5 comprising a sequence selected from the group consisting of SEQ ID NO:9, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:83, and fragments thereof; and

(b) sequencing said clone.

7. A method for obtaining a neurturin-persephin-GDNF family member growth factor comprising:

(a) isolating from a human genomic or cDNA template, a clone that, upon application of a polymerase
5 chain reaction method, produces an amplified polynucleotide having from about 30% to about 85% sequence identity to a region within a polynucleotide as set forth in a sequence selected from the group consisting of SEQ ID NO:9, SEQ ID NO:76, and SEQ ID
10 NO:81; and

192

(b) sequencing said clone.

8. The method according to claim 7 wherein the polymerase chain reaction method uses a degenerate primer obtained from a conserved region of any two of persephin, neurturin and GDNF and wherein the conserved region is as set forth in a sequence selected from the group consisting of SEQ ID NO:108, SEQ ID NO:113 and SEQ ID NO:114.

9. The method according to claim 8 wherein the conserved region degenerate primer has a sequence selected from the group consisting of SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, and SEQ ID NO:124.

10. An isolated and purified growth factor that is a neurturin-persephin-GDNF family member comprising a polypeptide having between about 30% and about 85% sequence identity with persephin, between about 30% and about 85% sequence identity with neurturin and between about 30% and about 85% sequence identity with GDNF.

11. The isolated and purified growth factor of claim 10 wherein said factor is comprised of a conserved region sequence of amino acids having at least a 62.5 percent sequence identity with SEQ ID NO:108 or at least a 62.5 percent sequence identity with SEQ ID NO:109 or at least a 50 percent sequence identity with SEQ ID NO:110.

12. The isolated and purified growth factor of claim 11 which is encoded by a nucleic acid identified and/or obtained by the polymerase chain reaction method utilizing a primer containing a nucleotide sequence selected from the group consisting of SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, and SEQ ID NO:124.

13. A pan-growth factor comprising a portion of the persephin polypeptide according to claim 1 and a

193

portion of at least one growth factor from the TGF-B
superfamily other than persephin.

14. The pan-growth factor according to claim 13
wherein the growth factor other than persephin is
neurturin and the pan-growth factor comprises SEQ ID
NO:141 or SEQ ID NO:146.

15. An isolated and purified nucleic acid
comprising a nucleotide sequence encoding a growth factor
of any one of claims 1-3 and 10-14.

16. The isolated and purified nucleic acid which
encodes a sequence selected from the group consisting of
SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82,
SEQ ID NO:83 and conservatively substituted variants
5   thereof.

17. The isolated and purified nucleic acid
comprising SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:174, SEQ
ID NO:175, and SEQ ID NO:176.

18. An isolated and purified nucleic acid sequence
which hybridizes to a nucleotide sequence complementary
to a nucleic acid sequence selected from the group
consisting of:

5   (a) a nucleic acid according to claim 15;
             (b) a nucleic acid sequence which encodes
             persephin according to claim 1; and
             (c) a nucleic acid sequence which encodes a
             polypeptide which cross-reacts with persephin as
10   defined in claim 1.

19. A vector comprising a recombinant DNA molecule
comprising expression regulatory elements operably linked
to a nucleic acid sequence according to claim 15.

20. A host cell transformed with the vector of
claim 19.

21. The host cell according to claim 20 wherein
said host cell is a mammalian cell, a bacterial cell or a
baculovirus expression system.

22. A recombinant DNA method comprising:

194

(a) subcloning a DNA sequence encoding a growth factor as defined in claim 1 into an expression vector which comprises regulatory elements needed to express the DNA sequence;

(b) transforming a host cell with said expression vector;

(c) growing the host cell in a host cell culture; and

(d) harvesting the growth factor and/or the DNA from the host cell culture.

23. Isolated and purified antibodies which are capable of reacting with a growth factor as defined in claim 1 or an epitope thereof.

24. A method for preventing or treating cellular degeneration or insufficiency in an individual comprising administering to the individual a therapeutically effective amount of a growth factor as defined in any of claims 1-3 and 10-14 or a DNA molecule encoding said factor.

25. The method of claim 24 wherein the cellular degeneration or insufficiency is neuronal degeneration resulting from a condition selected from the group consisting of peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, Ischemic stroke, acute brain injury, acute spinal cord injury, nervous system tumors, multiple sclerosis, and infection.

26. The method of claim 24 wherein the cellular degeneration or insufficiency comprises hematopoietic cell degeneration or insufficiency resulting from a condition selected from the from the group consisting of eosinopenia, basopenia, lymphopenia, monocytopenia, neutropenia, anemias, thrombocytopenia, and stem-cell insufficiencies therefor.

27. The method of claim 24 wherein the cellular degeneration or insufficiency comprises cardiac muscle

195

degeneration or insuffici ncy resulting from
cardiomyopathy or congestive heart failure.

28. A method for preventing or treating cellular
degeneration or insufficiency in an individual comprising
implanting into the individual cells that express a
growth factor as defined in any of claims 1-3 and 10-14.

29. A method for detecting the presence of a
growth factor in a sample from a patient comprising
reacting antibodies according to claim 23 with a growth
factor present in the sample and detecting a binding of
5 the antibodies with the growth factor.

30. A kit for detecting the presence of a growth
factor in a sample from a patient comprising antibodies
according to claim 23 which are capable of detectably
reacting with said growth factor packaged in a container.

31. A method for detecting the presence of a
growth factor in a sample from a patient comprising
detecting and/or quantitating the presence in the sample
of mRNA encoding a growth factor of claims 1.

32. A method for detecting persephin gene
alterations comprising detecting the presence of a non-
intact persephin gene in a cell wherein presence of the
non-intact gene indicates the presence of gene
5 alterations.

33. A method for promoting the growth and/or
differentiation of a cell in a culture medium comprising
adding to the culture medium a growth factor as defined
in claim 1.

34. An isolated and purified persephin antisense
polynucleotide comprising a sequence complementary to a
nucleic acid sequence of claim 15 and capable of
hybridizing to a naturally-occurring DNA or mRNA
5 polynucleotide sequence encoding persephin to prevent
transcription and/or translation of an encoded persephin
polypeptide.

196

35. A method for treating a disease condition
mediated by expression of persephin by a population of
cells comprising administering to said cells an
inhibitory effective amount of the antisense
5 polynucleotide of claim 34.
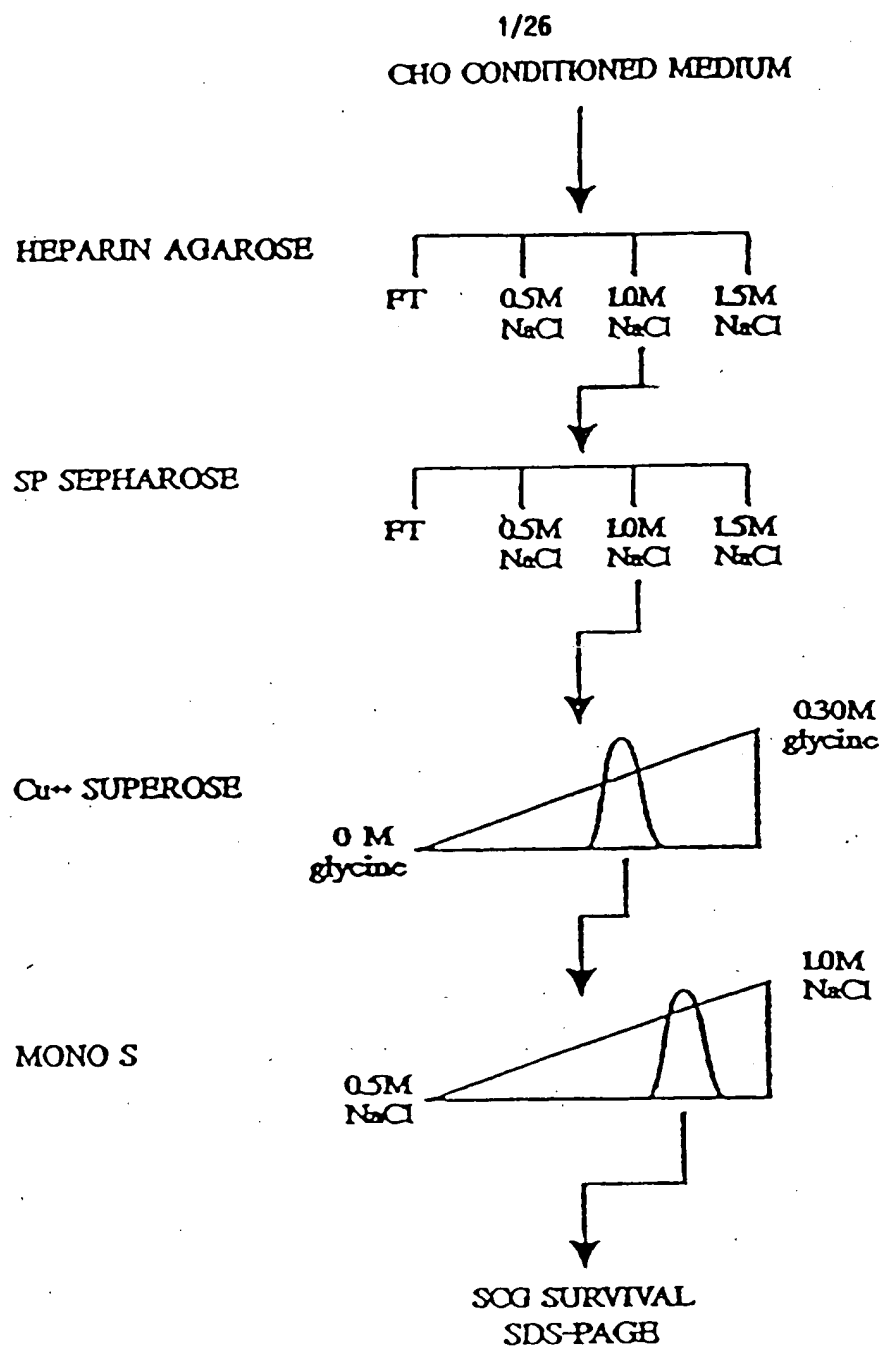
Figure 1
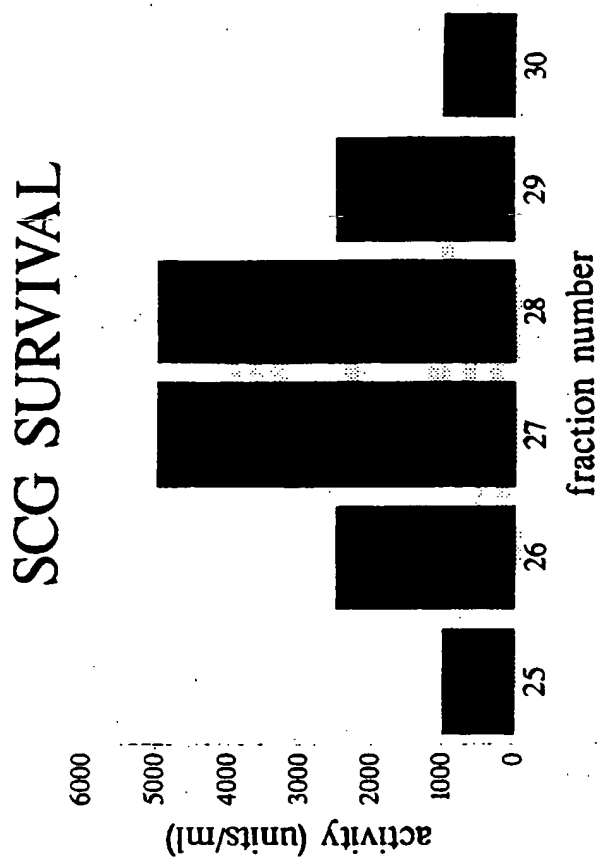
Figure 2a

Figure 2b

Figure 3a NGF
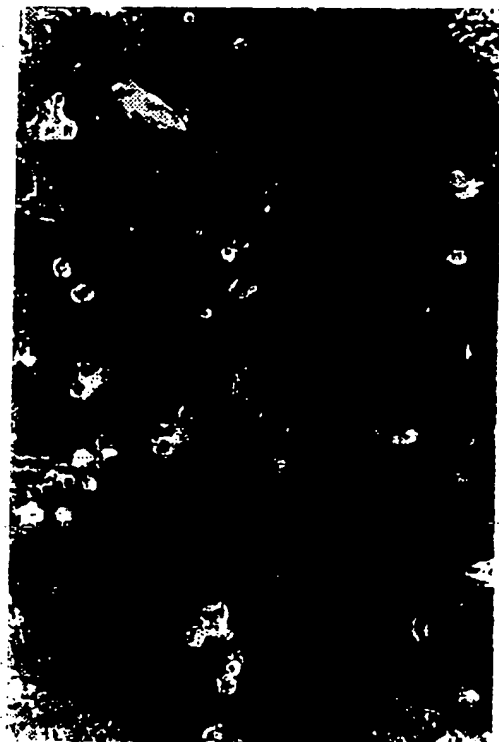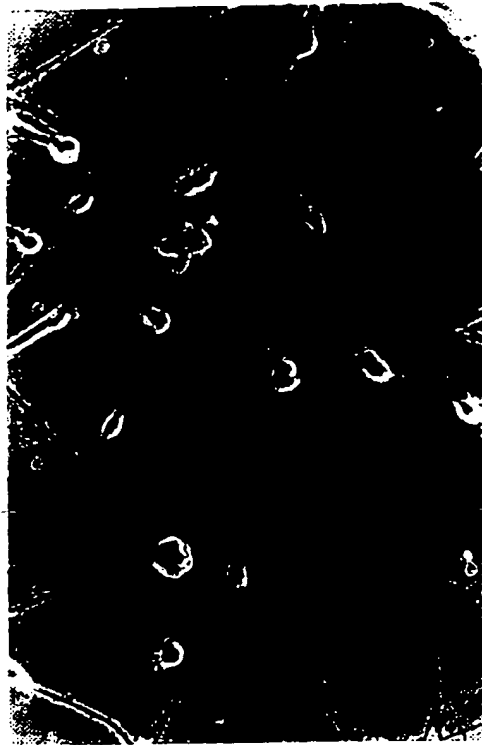
Figure 3b Anti-NGF

Figure 3c Anti-NGF
+
Neurturin
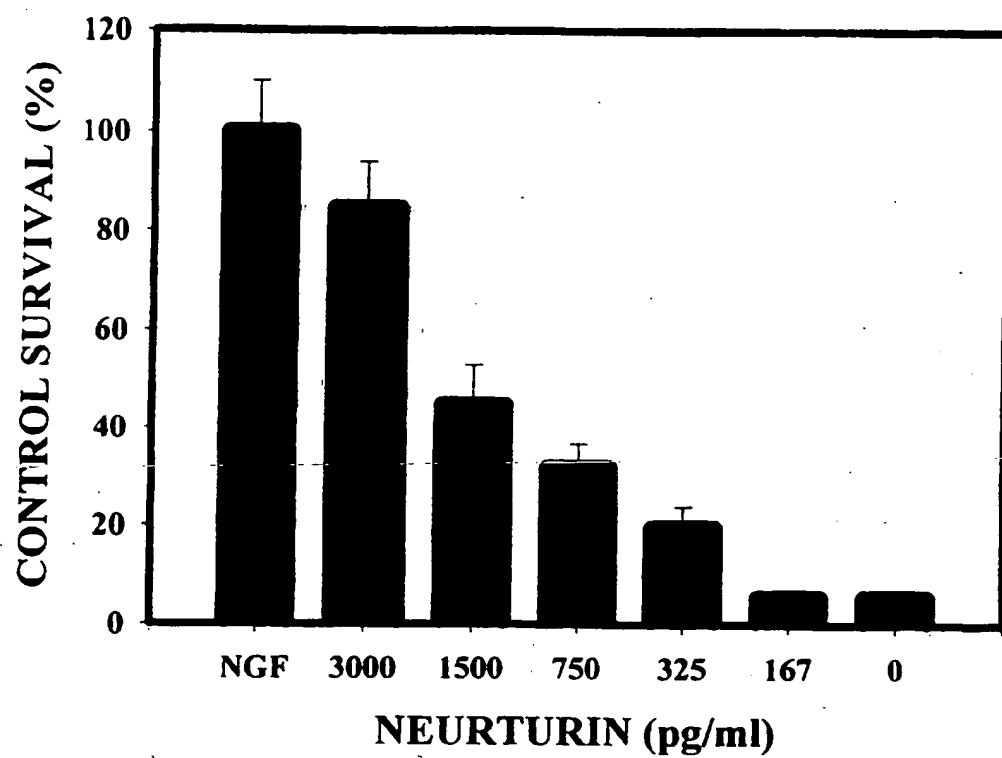
Figure 4

8/26

```
  1   S P D K Q M A V L P R R E R N R Q A A A A N P E N S R G K G   hGDNF
  1   S P D K Q A A A L P R R E R N R Q A A A A S P E N S R G K G   mGDNF
  1   S P D K Q A A A L P R R E R N R Q A A A A S P E N S R G K G   rGDNF
  1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   hNTN
  1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   mNTN

 31   R R G Q R G K N R G C V L T A I H L N V T D L G L G Y E T K   hGDNF
 31   R R G Q R G K N R G C V L T A I H L N V T D L G L G Y E T K   mGDNF
 31   R R G Q R G K N R G C V L T A I H L N V T D L G L G Y E T K   rGDNF
  1   - - - A R L G A R P C G L R E L E V R V S E L G L G Y A S D   hNTN
  1   - - - - - P G A R P C G L R E L E V R V S E L G L G Y T S D   mNTN

 61   E E L I F R Y C S G S C D A A E T T Y D K I L K N L S R N R   hGDNF
 61   E E L I F R Y C S G S C E S A E T M Y D K I L K N L S R S R   mGDNF
 61   E E L I F R Y C S G S C E A A E T M Y D K I L K N L S R S R   rGDNF
 28   E T V L F R Y C A G A C E A A A R V Y D L G L R R L R Q R R   hNTN
 26   E T V L F R Y C A G A C E A A I R I Y D L G L R R L R Q R R   mNTN

 91   R L V S D K V - G Q A C C R P I A F D D D L S F L D D N L V   hGDNF
 91   R L T S D K V - G Q A C C R P V A F D D D L S F L D D N L V   mGDNF
 91   R L T S D K V - G Q A C C R P V A F D D D L S F L D D S L V   rGDNF
 58   R L R R E R V R A Q P C C R P T A Y E D E V S F L D A H S R   hNTN
 56   R V R R E R A R A H P C C R P T A Y E D E V S F L D V H S R   mNTN

120   Y H I L R K H S A K R C G C I .                               hGDNF
120   Y H I L R K H S A K R C G C I .                               mGDNF
120   Y H I L R K H S A K R C G C I .                               rGDNF
 88   Y H T V H E L S A R E C A C V .                               hNTN
 86   Y H T L Q E L S A R E C A C V .                               mNTN
```

Figure 5

RT-PCR Analysis of Neurturin and GDNF
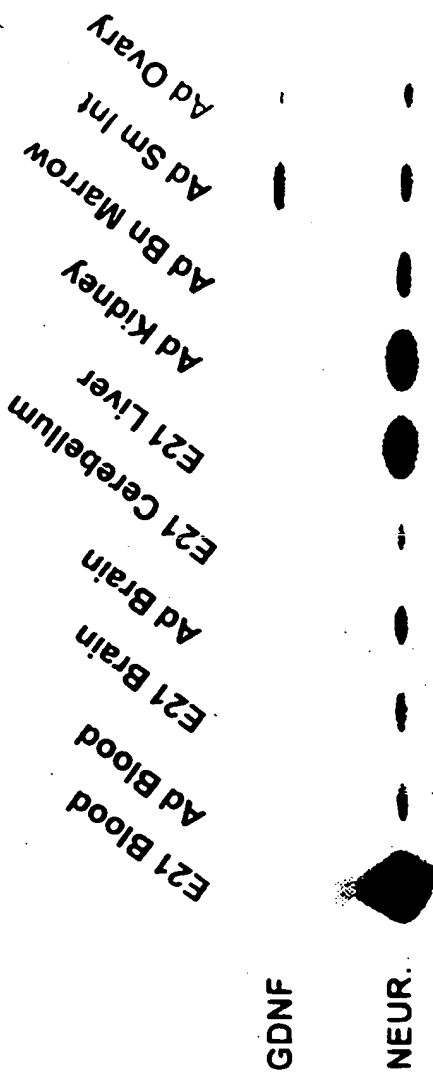
Figure 6

ATGCAGCGCTGGAAGGCGGCGGCCTTGGCCTCAGTGCTCTGCAGCTCCGTGCTGTCCATC 60
Met Gln Arg Trp Lys Ala Ala Ala Leu Ala Ser Val Leu Cys Ser Ser Val Leu Ser Ile

TGGATGTGTCGAGAGGGCCTGCTTCTCAGCCACCGCCTCGGACCTGCGCTGGTCCCCCTG 120
Trp Met Cys Arg Glu Gly Leu Leu Leu Ser His Arg Leu Gly Pro Ala Leu Val Pro Leu

CACCGCCTGCCTCGAACCCTGGACGCCCGGATTGCCCGCCTGGCCCAGTACCGTGCACTC 180
His Arg Leu Pro Arg Thr Leu Asp Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu

CTGCAGGGGGCCCCGGATGCGATGGAGCTGCGCGAGCTGACGCCCTGGGCTGGGCGGCCC 240
Leu Gln Gly Ala Pro Asp Ala Met Glu Leu Arg Glu Leu Thr Pro Trp Ala Gly Arg Pro

CCAGGTCCGCGCCGTCGGGCGGGGCCCCGGCGGCGGCGCGCGCGTGCGCGGTTGGGGGGCG 300
Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Arg Ala Arg Ala Arg Leu Gly Ala

CGGCCTTGCGGGCTGCGCGAGCTGGAGGTGCGCGTGAGCGAGCTGGGCCTGGGCTACGCG 360
Arg Pro Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly Tyr Ala

TCCGACGAGACGGTGCTGTTCCGCTACTGCGCAGGCGCCTGCGAGGCTGCCGCGCGCGTC 420
Ser Asp Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys Glu Ala Ala Ala Arg Val

TACGACCTCGGGCTGCGACGACTGCGCCAGCGGCGGCGCCTGCGGCGGGAGCGGGTGCGC 480
Tyr Asp Leu Gly Leu Arg Arg Leu Arg Gln Arg Arg Arg Leu Arg Arg Glu Arg Val Arg

GCGCAGCCCTGCTGCCGCCCGACGGCCTACGAGGACGAGGTGTCCTTCCTGGACGCGCAC 540
Ala Gln Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Ala His

AGCCGCTACCACACGGTGCACGAGCTGTCGGCGCGCGAGTGCGCCTGCGTGTGA 594
Ser Arg Tyr His Thr Val His Glu Leu Ser Ala Arg Glu Cys Ala Cys Val •

Figure 7

ATGAGGCGCTGGAAGGCAGCGGCCCTGGTGTCGCTCATCTGCAGCTCCCTGCTATCTGTC 60
Met Arg Arg Trp Lys Ala Ala Ala Leu Val Ser Leu Ile Cys Ser Ser Leu Leu Ser Val

TGGATGTGCCAGGAGGGTCTGCTCTTGGGCCACCGCCTGGGACCCGCGCTTGCCCCGCTA 120
Trp Met Cys Gln Glu Gly Leu Leu Leu Gly His Arg Leu Gly Pro Ala Leu Ala Pro Leu

CGACGCCCTCCACGCACCCTGGACGCCCGCATCGCCCGCCTGGCCCAGTATCGCGCTCTG 180
Arg Arg Pro Pro Arg Thr Leu Asp Ala Arg Ile Ala Arg Leu Ala Gln Tyr Arg Ala Leu

CTCCAGGGCGCCCCCGACGCGGTGGAGCTTCGAGAACTTTCTCCCTGGGCTGCCCGCATC 240
Leu Gln Gly Ala Pro Asp Ala Val Glu Leu Arg Glu Leu Ser Pro Trp Ala Ala Arg Ile

CCGGGACCGCGCCGTCGAGCGGGTCCCCGGCGTCGGCGGGCGCGGCCGGGGGCTCGGCCT 300
Pro Gly Pro Arg Arg Arg Ala Gly Pro Arg Arg Arg Arg Ala Arg Pro Gly Ala Arg Pro

TGTGGGCTGCGCGAGCTCGAGGTGCGCGTGAGCGAGCTGGGCCTGGGCTACACGTCGGAT 360
Cys Gly Leu Arg Glu Leu Glu Val Arg Val Ser Glu Leu Gly Leu Gly Tyr Thr Ser Asp

GAGACCGTGCTGTTCCGCTACTGCGCAGGCGCGTGCGAGGCGGCCATCCGCATCTACGAC 420
Glu Thr Val Leu Phe Arg Tyr Cys Ala Gly Ala Cys Glu Ala Ala Ile Arg Ile Tyr Asp

CTGGGCCTTCGGCGCCTGCGCCAGCGGAGGCGCGTGCGCAGAGAGCGGGCGCGGGCGCAC 480
Leu Gly Leu Arg Arg Leu Arg Gln Arg Arg Arg Val Arg Arg Glu Arg Ala Arg Ala His

CCGTGTTGTCGCCCGACGGCCTATGAGGACGAGGTGTCCTTCCTGGACGTGCACAGCCGC 540
Pro Cys Cys Arg Pro Thr Ala Tyr Glu Asp Glu Val Ser Phe Leu Asp Val His Ser Arg

TACCACACGCTGCAAGAGCTGTCGGCGCGGGAGTGCGCGTGCGTGTGA 588
Tyr His Thr Leu Gln Glu Leu Ser Ala Arg Glu Cys Ala Cys Val •

Figure 8

GGAGGGAGAGCGCGCGGTGGTTTCGTCCGTGTGCCCCGCGCCCGGCGC   -301

TCCTCGCGTGGCCCCGCGTCCTGAGCGCGCTCCAGCCTCCCACGCGCGCC   -251

ACCCCGGGGTTCACTGAGCCCGGCGAGCCCGGGGAAGACAGAGAAAGAGA   -201

GGCCAGGGGGGGAACCCCATGGCCCGGCCCGTGTCCCGCACCCTGTGCGG   -151

TGGCCTCCTCCGGCACGGGGTCCCCGGGTCGCCTCCGGTCCCCGCGATCC   -101

GGATGGCGCACGCAGTGGCTGGGGCCGGGCCGGGCTCGGGTGGTCGGAGG   -51

AGTCACCACTGACCGGGTCATCTGGAGCCCGTGGCAGGCCGAGGCCCAGG   -1

ATGAGGCGCTGGAAGGCAGCGGCCCTGGTGTCGCTCATCTGCAGCTCCCT   50

GCTATCTGTCTGGATGTGCCAGGAGGGTCTGCTCTTGGGCCACCGCCTGG   100

GACCCGCGCTTGCCCCGCTACGACGCCCTCCACGCACCCTGGACGCCCGC   150

ATCGCCCGCCTGGCCCAGTATCGCGCTCTGCTCCAGGGCGCCCCCGACGC   200

GGTGGAGCTTCGAGAACTTTCTCCCTGGGCTGCCCGCATCCCGGGACCGC   250

GCCGTCGAGCGGGTCCCCGGCGTCGGCGGGCGCGGCCGGGGGCTCGGCCT   300

TGTGGGCTGCGCGAGCTCGAGGTGCGCGTGAGCGAGCTGGGCCTGGGCTA   350

CACGTCGGATGAGACCGTGCTGTTCCGCTACTGCGCAGGCGCGTGCGAGG   400

CGGCCATCCGCATCTACGACCTGGGCCTTCGGCGCCTGCGCCAGCGGAGG   450

CGCGTGCGCAGAGAGCGGGCGCGGGCGCACCCGTGTTGTCGCCCGACGGC   500

CTATGAGGACGAGGTGTCCTTCCTGGACGTGCACAGCCGCTACCACACGC   550

TGCAAGAGCTGTCGGCGCGGGAGTGCGCGTGCGTGTGATGCTACCTCACG   600

CCCCCCGACCTGCGAAAGGGCCCTCCCTGCCGACCCTCGCTGAGAACTGA   650

CTTCACATAAAGTGTGGGAACTCCC   675
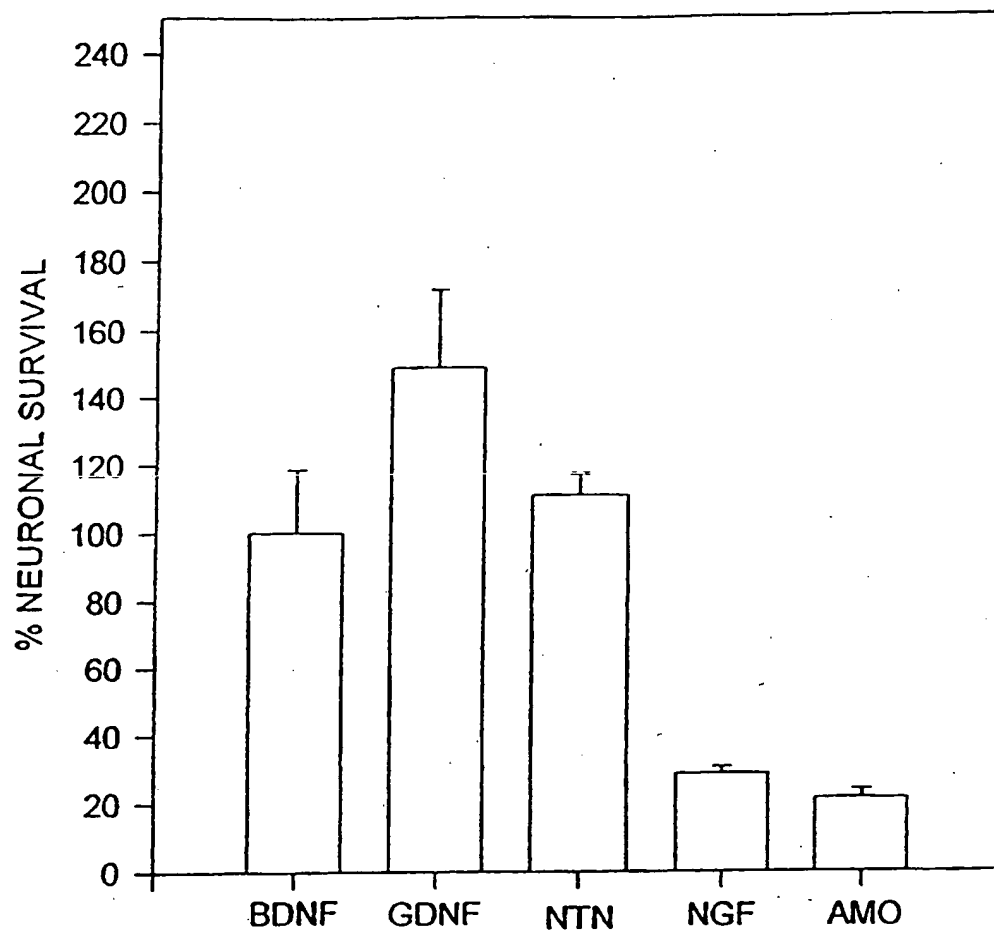
Figure 9

Figure 10

GAGGGACCTGGACGCCCCATCAGGGTAAGAATTCCTGGGGGCCTCCCGACTCCCCAATTC   60
Glu Gly Pro Gly Arg Pro Ile Arg Val Arg Ile Pro Gly Gly Leu Pro Thr Pro Gln Phe   20

CTTCTCTCAAAGCCCTCACTTTGCCTTACAATCCTACTCTACCTTGCACTAGGTAACAAC   120
Leu Leu Ser Lys Pro Ser Leu Cys Leu Thr Ile Leu Leu Tyr Leu Ala Leu Gly Asn Asn   40

CATGTCCGTCTTCCAAGAGCCTTGGCTGGTTCATGCCGACTGTGGAGCCTGACCCTACCA   180
His Val Arg Leu Pro Arg Ala Leu Ala Gly Ser Cys Arg Leu Trp Ser Leu Thr Leu Pro   60

GTGGCTGAGCTGGGCCTGGGCTATGCCTCGGAGGAGAAGGTCATCTTCCGATACTGTGCT   240
Val Ala Glu Leu Gly Leu Gly Tyr Ala Ser Glu Glu Lys Val Ile Phe Arg Tyr Cys Ala   80

GGCAGCTGTCCCCAAGAGGCCCGTACCCAGCACAGTCTGGTACTGGCCCGGCTTCGAGGG   300
Gly Ser Cys Pro Gln Glu Ala Arg Thr Gln His Ser Leu Val Leu Ala Arg Leu Arg Gly   100

CGGGGTCGAGCCCATGGCCGACCCTGCTGCCAGCCCACCAGCTATGCTGATGTGACCTTC   360
Arg Gly Arg Ala His Gly Arg Pro Cys Cys Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe   120

CTTGATGATCAGCACCATTGGCAGCAGCTGCCTCAGCTCTCAGCTGCAGCTTGTGGCTGT   420
Leu Asp Asp Gln His His Trp Gln Gln Leu Pro Gln Leu Ser Ala Ala Ala Cys Gly Cys   140

GGTGGCTGAAGGAGGCCAGTCTGGTGTCTCAGAATCACAAGCATGAGACAGGCTGGGCTT   480
Gly Gly   142

TGAAAGGCTCAGGTGACATTACTAGAAATTTGCATAGGTAAAGATAAGAAGGGAAAGGAC   540

CAGG   544

Figure 11

Figure 12

CCTCAGAGGAGAAGATTATCTTCCGATACTGTGCTGGCAGCTGTCCCCAAGAGGTCCGTACC   62
Ser Glu Glu Lys Ile Ile Phe Arg Tyr Cys Ala Gly Ser Cys Pro Gln Glu Val Arg Thr   20

CAGCACAGTCTGGTGCTGGCCCGTCTTCGAGGGCAGGGTCGAGCTCATGGCAGACCTTGC   122
Gln His Ser Leu Val Leu Ala Arg Leu Arg Gly Gln Gly Arg Ala His Gly Arg Pro Cys   40

TGCCAGCCCACCAGCTATGCTGATGTGACCTTCCTTGATGACCACCACCATTGGCAGCAG   182
Cys Gln Pro Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp Asp His His His Trp Gln Gln   60

CTGCCTCAGCTCTCAGCCGCAGCTTGTGGCTGTGGTGGCTGAAGGCGGCCAGCCTGGTCT   242
Leu Pro Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly   73

CTCAGAATCACAAGCAAGAGGCAGCCTTTGAAAGGCTCAGGTGACGTTATTAGAAACTTG   302

CATAGGAGAAGATTAAGAAGAGAAAGGGGACCTG   336

Figure 13

```
TGCCGGCTGTGGAGCCTGACCCTACCAGTGGCTGAGCTTGGCCTGGGCTATGCCTCAGAG        60
Cys Arg Leu Trp Ser Leu Thr Leu Pro Val Ala Glu Leu Gly Leu Gly Tyr Ala Ser Glu    20


GAGAAGATTATCTTCCGATACTGTGCTGGCAGCTGTCCCCAAGAGGTCCGTACCCAGCAC      120
Glu Lys Ile Ile Phe Arg Tyr Cys Ala Gly Ser Cys Pro Gln Glu Val Arg Thr Gln His    40


AGTCTGGTGCTGGCCCGTCTTCGAGGGCAGGGTCGAGCTCATGGCAGACCTTGCTGCCAG      180
Ser Leu Val Leu Ala Arg Leu Arg Gly Gln Gly Arg Ala His Gly Arg Pro Cys Cys Gln    60


CCCACCAGCTATGCTGATGTGACCTTCCTTGATGACCACCACCATTGGCAGCAGCTGCCT      240
Pro Thr Ser Tyr Ala Asp Val Thr Phe Leu Asp Asp His His His Trp Gln Gln Leu Pro    80


CAGCTCTCAGCCGCAGCTTGTGGCTGTGGTGGCTGAAGGCGGCCAGCCTGGTCTCTCAGA      300
Gln Leu Ser Ala Ala Ala Cys Gly Cys Gly Gly                                        91


ATCACAAGCAAGAGGCAGCCTTTGAAAGGCTCAGGTGACGTTATTAGAAACTTGCATAGG      360


AGAAGATTAAGAAGAGAAAGGGGACCTGATT                                                   391
```

Figure 14

```
             10           20          30          40
 1  C V L T A I H L N V T D L G L G Y E T K E E L I F R Y C S G S C D A A E T T - Y  GDNF(human)
 1  C V L T A I H L N V T D L G L G Y E T K E E L I F R Y C S G S C E A A E T M - Y  GDNF(rat)
 1  C V L T A I H L N V T D L G L G Y E T K E E L I F R Y C S G S C E S A E T M - Y  GDNF(mouse)
 1  C G L R E L E V R V S E L G L G Y A S D E T V L F R Y C A G A C E A A A R V - Y  NTN(human)
 1  C G L R E L E V R V S E L G L G Y T S D E T V L F R Y C A G A C E A A I R I - Y  NTN(mouse)
 1  C R L W S L T L P V A E L G L G Y A S E E K I I F R Y C A G S C P Q E V R T Q H  PSP(rat)
 1  C R L W S L T L P V A E L G L G Y A S E E K V I F R Y C A G S C P Q E A R T Q H  PSP(mouse)
```

```
             50           60          70          80
40  D K I L K N L S R N R R L V S D K V - G Q A C C R P I A F D D D L S F L D D N L  GDNF(human)
40  D K I L K N L S R S R R L T S D K V - G Q A C C R P V A F D D D L S F L D D S L  GDNF(rat)
40  D K I L K N L S R S R R L T S D K V - G Q A C C R P V A F D D D L S F L D D N L  GDNF(mouse)
40  D L G L R R L R Q R R R L R R E R V R A Q P C C R P T A Y E D E V S F L D A H S  NTN(human)
40  D L G L R R L R Q R R R V R R E R A R A H P C C R P T A Y E D E V S F L D V H S  NTN(mouse)
41  S L V L A R L R - - - - - G Q G R A H G R P C C Q P T S Y A D - V T F L D D H H  PSP(rat)
41  S L V L A R L R - - - - - G R G R A H G R P C C Q P T S Y A D - V T F L D D Q H  PSP(mouse)
```

```
             90
79  V Y H I L R K H S A K R C G C .  GDNF(human)
79  V Y H I L R K H S A K R C G C .  GDNF(rat)
79  V Y H I L R K H S A K R C G C .  GDNF(mouse)
80  R Y H T V H E L S A R E C A C .  NTN(human)
80  R Y H T L Q E L S A R E C A C .  NTN(mouse)
75  H W Q Q L P Q L S A A A C G C .  PSP(rat)
75  H W Q Q L P Q L S A A A C G C .  PSP(mouse)
```

Figure 15

```
SEQ
ID    GROWTH
NO.   FACTOR   SEQUENCE

150   TGFB1    CCVRQLYIDFRKDLGWK-WIHEPKGYHANFCLGPCPYIWSLDT------QYSKVLALYNQHNPGASAA-PCCV--PQALEPLPIVYYVGRKPKV--EQLSNMIVRSCKCS
151   TGFB2    CCLRPLYIDFKRDLGWK-WIHEPKGYNANFCAGACPYLWSSDT------QHSRVLSLYNTINPEASAS-PCCV--SQDLEPLTILYYIGKTPKI--EQLSNMIVKSCKCS
152   TGFB3    CCVRPLYIDFRQDLGWK-WVHEPKGYYANFCSGPCPYLRSADT------THSTVLGLYNTLNPEASAS-PCCV--PQDLEPLTILYYVGRTPKV--EQLSNMVVKSCKCS
153   INHBA    CCKKQFFVSFK-DIGWNDWIIAPSGYHANYCEGECPSHIAG-TSGSSLSFHSTVINHYRMRGHSPFANLKSCCV--PTKLRPMSMLYYDDGQNII-KKDIQNMIVEECGCS
154   INHBB    CCRQQFFIDFR-LIGWNDWIIAPTGYYGNYCEGECPSCPAYLAG-VPGSASSFHTAVVNQYRMRGLNP-GTVNSCCI--PTKLSTMSMLYFDDEYNIV-KRDVPNMIVEECGCA
155   NODAL    CRRVKFQVDFN-LIGWGSWIIYPKQYNAYRCEGECPNPVGEEFHPT----NHAYIQSLLKRYQPHR-VPSTCCA--PVKTKPLSMLYVDNGR--VLLEHHKOMIVEECGCL
156   BMP2     CKRHPLYVDFS-DVGWNDWIVAPPGYHAFYCHGECPFPLADHLNST----NHAIVQTLVNSVNS-K-IPKACCV--PTELSAISMLYLDENEKVVLK-NYQDMVVEGGCGCR
157   BMP4     CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGDCPFPLADHLNST----NHAIVQTLVNSVNS-S-IPKACCV--PTELSAISMLYLDEYDKVVLK-NYQEMVVEGCGCR
158   DPP      CRRHSLYVDFS-DVGWDDWIVAPLGYDAYYCHGKCPFPLADHFNST----NHAVVQTLVNNMNPGK-VPKACCV--PTQLDSVAMLYLNDQSTVVLK-NYQEMTVVGGGCR
159   BMP5     CKKHELYVSFR-DLGWQDWIIAPEGYAAFYCDGECPFPLNAHMNAT----NHAIVQTLVHLMFPDH-VPKPCCA--PTKLNAISVLYFDDSSNVILK-KYRNMVVRSCGCH
160   BMP6     CRKHELYVSFQ-DLGWQDWIIAPKGYAANYCDGECSFPLNAHMNAT----NHAIVQTLVHLMNPEY-VPKPCCA--PTKLNAISVLYFDDNSNVILK-KYRNMVVRACGCH
161   BMP7     CKKHELYVSFR-DLGWQDWIIAPEGYAAYYCEGECAFPLNSYMNAT----NHAIVQTLVHFINPET-VPKPCCA--PTQLNAISVLYFDDSSNVILK-KYRNMVVRACGCH
162   BMP8     CRRHELYVSFQ-DLGWLDWVIAPQGYSAYYCEGECSFPLDSCMNAT----NHAILQSLVHLMKPNA-VPKACCA--PTKLSATSVLYQDSSNNVILR-KHRNMVVKACGCH
163   60A      CQMTLYIDFK-DLGWHDWIIAPEGYGAFYCSGECNFPLNAHMNAT----NHAIVQTLVHLLEPKK-VPKPCCA--PTRLGALPVLYHLNDENVNLK-KYRNMIVKSCGCH
164   BMP3     CARRYLKVDFA-DIGWSEWIISPKSFDAYYCSGACQFPMPKSLKPS----NHATIQSIVRAVGVVPGIPEPCCV--PEKMSSLSILFFDENKNVVLKV-YPNMTVESCACR
165   VG1      CKKRHLYVEFK-DVGWQNWVIAPQGYMANYCYGECPYPLTEILNGS----NHAILQTLVHSIEPED-IPLPCCV--PTKMSPISMLFYDNNDNNVLR-HYENMAVDECGCR
166   GDF1     CRARRLYVSFR-EVGWHRWVIAPRGFLANYCQGQCALPVALSGSGGPPALNHAVLRALMHAAAPGA-ADLPCCV--PARLSPISVLFFDNSDNVVLR-QYEDMVVDECGCR
167   GDF3     CHRHQLFINFQ-DLGWHKWVIAPKGFMANYCHGECPFSMTTYLNSS----NYAFMQALMHMADP-K-VPKAVCV--PTKLSPISMLYQDSDKNVILR-HYEDMVVDECGCG
168   DORSLN   CRRTSLHVNFK-EIGWDSWIIAPKDYEAFECKGGCFFPLTDNVTPT----KHAIVQTLVHLQNPKK-ASKACCV--PTKLDAISILYKDDAGVPTLIYNYEGMKVAECGCR
169   INHα     CHRVALNISF-QELGWERWIVYPPSFIFHYCHGGCGGLHIPPNNLSLPVPGAPPTPAQPYSLL-----PGAQPCCAALPGTMRPLHVRTTSDGGYSFKYETVPNLLTQHCACI
170   MIS      CALRELSVDLRAERS-----VLIPETYQANNCQGACGWPQSDR----NPRYGNHVVLLLKMQARGATLARPPCCV--PTAYT--GKLLISLSEERISAHVPNMVATEGCGR
171   GDF9     CELHDFSLSFS-QLKWDNWIVAPHSYNPSYCKGDCPSAVSHRYGSPV---HTMVQNMIYE-KLDPSVPSPSCV--PGKYSPLSVLTIEPDGSIAYK-EYEDMATSCTCR
172   GDNF     CVLTAIHLNVT-DLGLG--YETKEELIFRYCSGSCD-AAETTYDKILKNLSRN-----RRLVSDKV-GQACCRPIAFD-DDLSFL------DDNLVHILRKHSAKRCGCI
173   NTN      CGLRELEVRVS-ELGLG--YASDETVLFRYCAGACE-AAARVYDLGLRRLRQR-----RRLRRERVRAQPCCRPTAYE-DEVSFL------DAHSRYHTVHELSARECACV
```

FIGURE 16

```
ATGGCTGCAGGAAGACTTCGGATCCTGTGTCTGCTCCTGTCCTTGCTCCGGTCGGCTGGTCCTTCAAGAGGCTTCTGTGTGGCAGATA   100
 M  A  A  G  R  L  R  I  L  C  L  L  L  S  L  L  P  S  L  G  W  V  L  D  L  Q  E  A  S  V  A  D   33
  W  L  Q  E  D  F  G  S  C  V  C  C  S  C  P  C  T  R  A  S  A  G  S  L  I  F  K  R  L  W  Q  I   33

AGCTCTCATTTGGGAAAGATGGCAGAGAGACTAGAGGGACCTGGACGCCCCATCAGGGTAAGAATTCCTGGGGGCCTCCCGACTCCCAATTCCTTCTCAA   200
 K  L  S  F  G  K  M  A  E  T  R  G  T  W  T  P  H  Q  G  K  N  S  W  G  P  P  D  S  P  I  P  S  L  K   67
  S  S  H  L  G  R  W  Q  R  L  E  G  P  G  R  P  I  R  V  R  I  P  G  G  L  P  T  P  Q  F  L  L  S   66

AGCCCTCACTTTGCCTTACAATCTACTTGCACTAGGTAACAACCATGTCCGTCTTCCAAGAGCCTTGGCGTCGGTTCATGCCGACTGTGGAGCCT   300
 A  L  T  L  P  Y  N  P  T  L  P  C  T  R                                                          81
  K  P  S  L  C  L  T  L  L  Y  L  A  L  G  N  N  H  V  R  L  P  R  A  L  A  G  S  C  R  L  W  S  L  100

GACCCTACCAGTGGCTGGGCTGAGCTGGGCCTGGGCTATGCCTCCGGAGGAGAAGGTCATCTTCCGATACTGTGCTGGCAGCTGTCCCCAAGAGGCCCGTACCCAG   400
 T  L  P  V  A  E  L  G  L  G  Y  A  S  E  E  K  V  I  F  R  Y  C  A  G  S  C  P  Q  E  A  R  T  Q  133

CACAGTCTGGTACTGGCCCGGCTTCGAGGGCGGGGTCGAGCCCATGGCCGACCCTGCTGCAGCCACCAGCTATGCTGATGTGACCTTCCTTGATGATC   500
 H  S  L  V  L  A  R  L  R  G  R  G  R  A  H  G  R  P  C  C  Q  P  T  S  Y  A  D  V  T  F  L  D  D  166

AGCACCATTGGCAGCAGCTGCCTCAGCTCTCAGCTGCCAGCTTGTGGCTGTGGCTGGGCTGA   559
 Q  H  H  W  Q  Q  L  P  Q  L  S  A  A  A  C  G  G  G              185
```

Figure 17

```
ATGGCTGCAGGAAGACTTCGGGATCTTGTTTCTGCTGCTCCTTGCTGCTGGGCCTTGGCTGGGGTCCTTGATCTTCAAGAGGCTCCTGCGGGCAGATG   100
 M  A  G  R  L  R  I  L  F  L  L  L  S  L  H  L  G  L  G  W  V  L  D  L  Q  E  A  P  A  A  D          33
  W  L  Q  E  D  F  G  S  C  F  C  C  S  C  P  C  T  W  A  L  L  A  G  S  L  I  F  K  R  L  R  Q  M    33

AGCTCTCATCTGGGAAAATGGCAGAGACTGGAAGCCCCATCAGGGTAAGAATTCTTGGGGCCTCCTAACTCTACAGTTCTTCCTCTCAA   200
 E  L  S  S  G  K  M  A  E  T  G  R  T  W  K  P  H  Q  G  K  N  S  W  G  P  P  N  S  T  V  L  P  L  K   67
  S  S  H  L  G  K  W  Q  R  L  E  G  P  G  S  P  I  R  V  R  I  L  G  G  L  L  T  L  Q  F  F  L  S    66

AGCCCTCACTTGCCTCACAATCCTATTCTACCTTGCACTAGTAACAACATGTCCGCCTTCCAAGAGCCTTACCTGGTTTGTCGCGGCTGTGGAGCCT   300
 A  L  T  L  P  H  N  P  I  L  P  C  T  R                                                            81
  K  P  S  L  C  L  T  I  L  F  Y  L  A  L  G  N  N  N  V  R  L  P  R  A  L  P  G  L  C  R  L  W  S  L  100

GACCCTACCAGTGGCTGAGCTTGGCCTGGGCTATGCCTCAGAGGAGAAGATTATCTTCCGATACTGTGCTGGCAGCTGTCCCCAAGAGGTCCGTACCCAG   400
 T  L  P  V  A  E  L  G  L  G  Y  A  S  E  K  I  I  F  R  Y  C  A  G  S  C  P  Q  E  V  R  T  Q        133

CACAGTCTCGGTGCTGGGCCCGTCTCTTCCGAGGGCAGGGTCGAGCTCATGGCAGACCTTGCTGCCAGCCCACCAGCTATGCTGATGTGCACCTTCCTTGATGACC   500
 H  S  L  V  L  A  R  L  R  G  Q  G  R  A  H  G  R  P  C  C  Q  P  T  S  Y  A  D  V  T  F  L  D  D      166

ACCACCATTGGCAGCAGCTGCCTCAGCTCTCAGCTCGCCAGCTTGTGGGCTGTGGGCTGA   559
 H  H  W  Q  Q  L  P  Q  L  S  A  A  A  C  G  G                 185
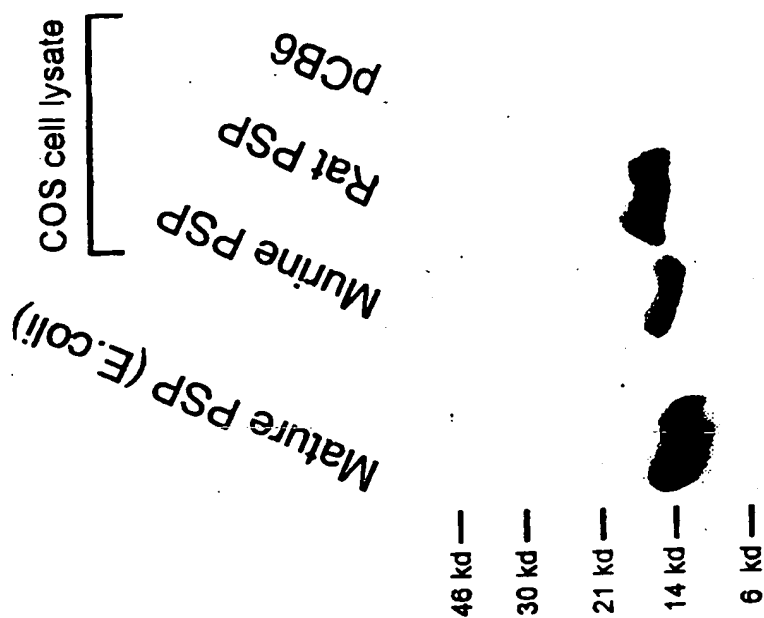```

Figure 18

Figure 19

23/26

# FIGURE 20A

## PSP/NTN (SEQ ID NO:137)

ALAGSCRLWSLTLPVAELGLGYASEEKVIFRYCAGSCPQEARTQHSLVLA          50
   ↓
RLRGRGRAHGRPCCRPTAYEDEVSFLDVHSRYHTLQELSARECACV            96


# FIGURE 20B

## NTN/PSP (SEQ ID NO:142)

PGARPCGLRELEVRVSELGLGYTSDETVLFRYCAGACEAAIRIYDLGLRR          50
   ↓
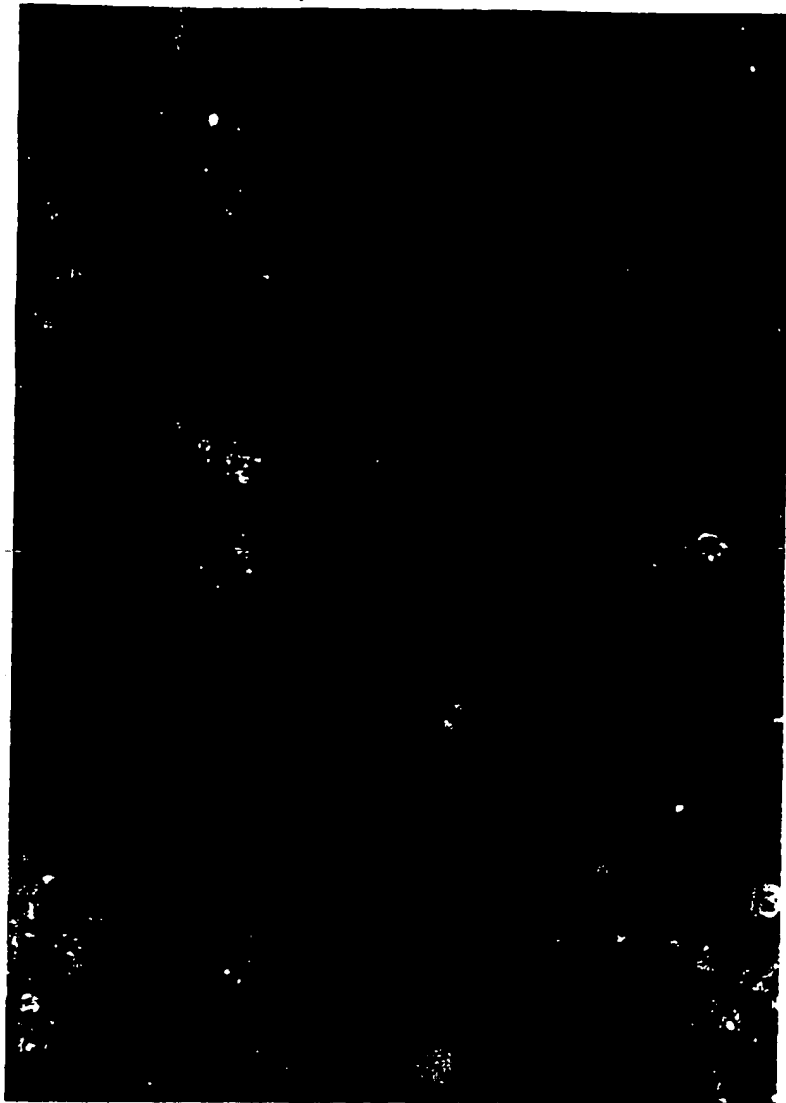LRQRRRVRRERARAHPCCQPTSYADVTFLDDQHHWQQLPQLSAAACGCGG         100

Figure 21a
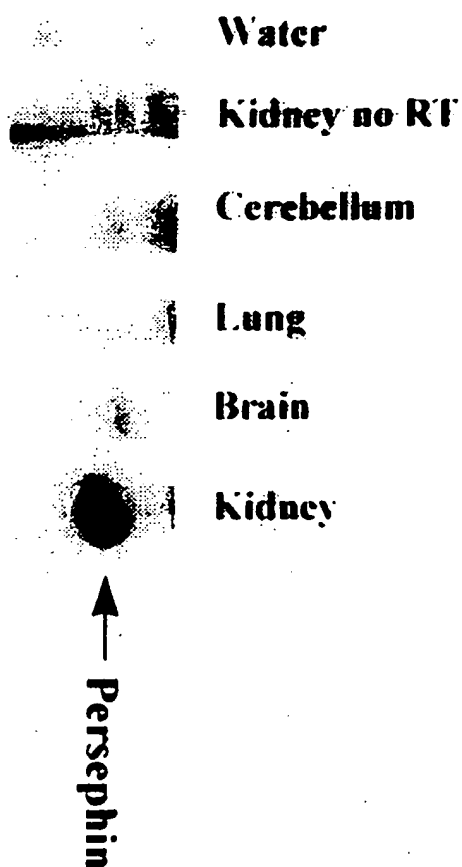
Figure 21b

Figure 22

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6)  :Please See Extra Sheet.

US CL  :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC.

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. :  530/399, 387.1; 435/69.4, 69.7, 325, 252.3, 320.1, 6, 7.1, 91.2; 536/23.51, 24.1, 24.31; 514/12, 44; 424/93.21

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, IntelliGenetics

search terms: persephin, GDNF, neurturin, glial derived neurotrophic factor#

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | LIN et al. GDNF: A Glial Cell Line-Derived Neurotrophic Factor for Midbrain Dopaminergic Neurons. Science. 21 May 1993. Vol. 260. pages 1130-1132. | 2-3, 5-12, 15-22, 34 |
| X | WATABE et al. Spontaneously Immortalized Adult Mouse Schwann Cells Secrete Autocrine and Paracrine Growth-Promoting Activities. Journal of Neuroscience Research. 01 June 1995. Vol 41. No. 2. pages 279-290. | 2-3, 5-12, 15-22, 34 |
| X | SPRINGER et al. CDNA Sequence and Differential mRNA Regulation of Two Forms of Glial Cell Line-Derived Neurotrophic Factor in Schwann Cells and Rat Skeletal Muscle. Experimental Neurology. January 1995. Vol. 131. No. 1. pages 47-52. | 2-3, 5-12, 15-22, 34 |

| X | Further documents are listed in the continuation of Box C. | | See patent family annex. |

\*        Special categories of cited documents:

"A"      document defining the general state of the art which is not considered to be of particular relevance

"E"      earlier document published on or after the international filing date

"L"      document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O"      document referring to an oral disclosure, use, exhibition or other means

"P"      document published prior to the international filing date but later than the priority date claimed

"T"      later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"      document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"      document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"      document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 03 JULY 1997 | 2 6 AUG 1997 |

| Name and mailing address of the ISA/US | Authorized Officer |
|---|---|
| Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 | ROBERT D. HAYES, PH.D. |
| Facsimile No. (703) 305-3230 | Telephone No. (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 95/17203 A1 (THE UNIVERSITY OF MEDICINE AND DENTISTRY OF NEW JERSEY) 29 June 1995, see entire document. | 2-3, 5-12, 15-22, 24-25, 34 |
| A | JACKOWSKI, A. Neural Injury Repair: Hope for the Future as Barriers to Effective CNS Regeneration Become Clearer. British Journal of Neurosurgery. 1995. Vol. 9. pages 303-317. | 1-35 |
| A | STULL et al. Antigene, Ribozyme and Aptamer Nucleic Acid Drugs: Progress and Prospects. Pharmaceutical Research. 1995. Vol. 12. No. 4. pages 465-483. | 24-28, 34-35 |
| A | SLOAN et al. The Immune Response to Intracerebral Neural Grafts. TINS. 1991. Vol. 14. No 8. pages 341-346. | 28 |
| X,P | KOTZBAUER et al. Neurturin, a Relative of Glial-Cell-Line-Derived Neurotrophic Factor. Nature. 05 December 1996. Vol. 384. No. 6608. pages 467-470. | 2-3, 5-12, 15-22, 34 |

| INTERNATIONAL SEARCH REPORT | International application No.<br>PCT/US97/03461 |
|---|---|

**Box I   Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely.

2. ☐ Claims Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II   Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

   Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**     ☐ The additional search fees were accompanied by the applicant's protest.
                          ☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)*

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07K 14/475, 16/26; C12N 15/18, 15/85, 15/63, 1/21; C07H 21/00; A61K 38/18, 48/00; C12Q 1/68; G01N 33/53;
C12P 19/34

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

530/399, 387.1; 435/69.4, 69.7, 325, 252.3, 320.1, 6, 7.1, 91.2; 536/23.51, 24.1, 24.31; 514/12, 44; 424/93.21

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single
inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search
fees must be paid.

Group I, claim(s) 1-5 and 10-22, drawn to growth factors comprising persephin, or pan-growth factors, pharmaceutical
compositions thereof, vectors, host cells, and method of producing these growth factors recombinantly.
Group II, claim(s) 6-9, drawn to a method of obtaining a neurturin-persephin-GDNF family member growth factor
using hybridization protocols, or PCR.
Group III, claim(s) 23, drawn to antibodies of growth factors that comprise persephin.
Group IV, claim(s) 24-27 and 33, drawn to methods of preventing/treating cellular degeneration comprising
administering to a patient a therapeutically effective amount of a growth factor comprising persephin, and in vitro
methods of promoting growth.
Group V, claim(s) 24-28, drawn to gene/cell therapy methods of preventing/treating cellular degeneration comprising
administering to a patient a DNA molecule encoding a growth factor comprising persephin, or implanting cells
expressing the growth factor.
Group VI, claim(s) 29-30, drawn to kits and methods of detecting a growth factor comprising persephin in a sample
comprising use of antibodies.
Group VII, claim(s) 31-32, drawn to methods of detecting a growth factor comprising persephin in a sample comprising
detecting nucleic acids.
Group VIII, claims 34-35, drawn to antisense persephin polynucleotides and gene therapy methods using the antisense
molecule to inhibit persephin expression in a population of cells.


The inventions listed as Groups I-VIII do not relate to a single inventive concept under PCT Rule 13.1 because, under
PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:
      Group I is directed to growth factors comprising persephin, or pan-growth factors, pharmaceutical
compositions thereof, vectors, host cells, and a method of producing these growth factors recombinantly, which is the
first appearing product, method of use and method of making the growth factors. The special technical feature is the
persephin growth factor protein. Groups II and IV-VIII are drawn to methods having different goals, method steps and
starting materials, which do not require each other for their practice and do not share the same or a corresponding
technical feature. Groups I, III, VI and VIII are drawn to structurally different products, which do not require each
other for their practice and do not share the same or a corresponding technical feature. Note that PCT Rule 13 does
not provide for multiple products or methods within a single application. Since the special technical feature of the
Group I invention is not present in the Group II-VIII claims, and the special technical features of the Group II-VIII
inventions are not present in the Group I claims, unity of invention is lacking.